Universidad Autónoma de Madrid

Universite d'Evry val d'Essonne

PhD Thesis

Gas-phase reactivity of Lewis-adducts and model biochemical systems. Quantum chemistry and molecular dynamics perspectives.

Author:
Ana Martín Sómer

Supervisors:
Manuel Yañez Montero
Riccardo Spezia

September 8, 2014

Resumen

La reactividad en fase gas, tambin conocida como *reactividad intrnseca*, tiene un gran interes ya que la ausencia de interacciones con un solvente puede resultar en una reactividad muy diferente, permitiendo adems un mejor concimiento de las propiedades moleculares. Con la aparicin en 1900 de nuevas tcnicas experimentales, concretamente tcnicas de ionizacin ms suaves, la qumica de iones en fase gas experiment un desarrollo significante con el consecuente cambio en cuanto a nuestra idea de reactividad qumica. El presente manuscrito est dividido en dos partes diferentes cada una de las cuales trata un aspecto distinto sobre la reactividad en fase gas.

La primera parte, Part I, trata sobre el estudio de la acidez intrnseca de una serie de bases de Lewis pertenecientes al grupo 13 de la tabla peridica. EL principal tema de estudio durante esta primera parte son los cambios que se producen en dichas acideces intrnsecas al formase un complejo de Lewis con distintos cidos de Lewis en los que el centro activo es un elemento del grupo 13 de la Tabla peridica. Asi, el origen del aumento de acidez observado al formarse el complejo es racionalizado a travs del uso de diferentes mtodos tericos. Se han llevado a cabo calculos DFT y *ab intio* de alto nivel con el fin de calcular los valores tericos para la acidez de las moleculas consideradas. Complementariamente, se usaron mtodos para el anlisis de la poblacin electrnica (AIM, NBO y ELF) con el objetivo de describir los cambios en la configuracin electrnica que se producen al formarse el complejo y que son responsables del observado aumento de acidez. Es importante destacar que parte de los resultados aqu presentados han sido corroborados mediante tcnicas experimentales.

La seguna parte, Part II, se centra en el estudio de la reactividad unimolecular de los iones formamida-M^{2+} (M = Ca, Sr). En este caso, existan trabajos experimentales disponibles relacionados con el estudio de la reacticvidad inducida por colisin (CID, en sus siglas en ingls) de dichos iones. A lo largo de la segunda parte de este manuscrito, se estudian y caracterizan los mecanismos de fragmentacin de ambos iones, empleando para ello distintas tcnicas tericas que son a la vez complementarias entre ellas. El primer paso en este estudio es la evaluacin de distintos mtodos tericos con el objetivo de encontrar el ms adecuado para llevar a cabo clculos fiables a la vez que se mantiene el coste computacional lo ms bajo posible. Por un lado, hemos realizado un estudio cintico del proceso de fragmentacin usando la teora estadstica conocida como RRKM, con el fin de describir la reactividad "lenta" (en una escala de tiempo $>$ fs). Por otro lado, realizamos simulaciones dinmicas que permiten describir la reactividad no estadstica en una escala corta de tiempo ($<$ 2.5 fs). De este modo, usando el descrito procedimiento

multi-escala, hemos sido capaces de racionalizar el origen de todos los productos observados en los espectros CID de los iones formamida-Ca^{2+} y formamida-Sr^{2+} obtenidos experimentalmente, as como las diferencias entre ellos.

En un cuarto capitulo se enumeran y describen brevemente los mtodos utilizados, tanto tericos como experimentales, a lo largo del trabajo presentado en este manuscrito.

Abstract

The so-called *intrinsic rectivity* (gas-phase reactivity) is of great importance since the absence of interaction with a solvent can result in very different reactivity patterns; allowing for a better understanding of molecular properties. With the advent in the 1900s of new experimental techniques, notably soft ionization methods such as electrospray ionization, the gas-phase ion chemistry has signigicantly developped in the last decades of the 1900s with a concomitant change in our view of chemical reactivity. The present manuscript is divide in two different parts each one dealing with different aspects of gas-phase reactivity.

Part I is concerned with the study of the intrinsic acidity of a series of group 15 Lewis base. The changes on the aforementioned intrinsic acidity as the Lewis bases form adducts with group 13 Lewis acid is the main subject of this part. Thus, the origin for the acidity enhancement observed upon adduct formation is rationalized by means of different theoretical methods. High-level DFT and *ab intio* calculations were perfomed in order to compute theoretical acidites of the molecules under survey. Complementary to this, population analysis techniques such as AIM, ELF, and NBO were used to analyze the changes on the electronic configurations of those molecules and therefore provide with an explanation to the observed acidities. It is worth to stress the fact that part of the results were as well confirmed by means of experimental measurements.

Part II focuses in unimolecular reactivity of molecular ions, namely, formamide-M^{2+} (M = Ca, Sr). In this case, experiments studying the Collision Induced Reactivity (CID) of these ions were already performed and through the second part of this manuscript the fragmentation mechanism of both ions are studied and characterized using different, but complementary, theoretical techniques. It is worth to mention that in a very first-step, an assessment of different methods to perform reliable electronic structure calculations while maintaining the lower possible computational cost. On the one hand, a kinetic study of the fragmentation process using the statistical theory, RRKM, to describe the long-time reactivity ($>$ fs). On the other hand, direct dynamics simulations are performed in order to describe the short-time ($<$ 2.5 fs) non-statistical reactivity. This multi-scale approach allowed us to account for all the products observed in the CID experimental spectra of formamide-M^{2+} (M = Ca, Sr), as well as the differences between them.

In the fourth chapter a summary of the experimental and theoretical procedures used to perform the work presented in this manuscript is provided.

Contents

Introduction

The work presented in this PhD thesis has been divided in two parts.

Part I of this document report's on the changes on the intrinsic acidity of a series of group 15 Lewis bases triggered by Lewis acid (BeH_2, BH_3, AlH_3, and GaH_3) complexation. This is done by means of high-level G4 *ab initio* calculations and Density Functional Theory (DFT) methods. The association with any of the Lewis acids considered leads to a significant acidity enhancement. In terms of the acidity constant, this increase can be as high as 40 orders of magnitude. To study the origin of this acidity enhancement the electron density rearrangements undergone by the Lewis acid and the Lewis base moieties upon adduct formation were analyzed by means of the complementary AIM, ELF, and NBO electronic density distribution analysis approaches. The dependence of the acidity enhancement observed on the nature of the Lewis acid, the active center of the Lewis base and the R substituent of the latter is also analyzed.

Part II focus on the modelization and characterization of the dynamical processes that occur in Collision Induced Dissociation (CID) experiments.

Aqui escribo la introduccion a parte de los otros apartados

Por que es interesante la quimica en fase gas? The development of gas-phase ion chemistry in the last decades of the 1900s led to a significant change in our view of chemical reactivity. The absence of solute-solvent and counterion interactions revealed the existence of reactivity trends which were very different from those usually accepted and obtained in condensed media. One point of interest was whether a result observed in the condensed phase carries over to the gas-phase, that is, is an intrinsic effect or a solvent effect. As a consequence, a great deal of effort was concentrated on determining intrinsic reactivities, in particular intrinsic basicities and acidities. [ref2-5 de Marcela]

In the first part the gas-phase acidities of series of molecules by means of electronic structure methods, i. e. a static aproach, are studied. This approach is perfectly suitable for processes thermodynamically controlled where the final products correspond to the most stable ones. However, there are

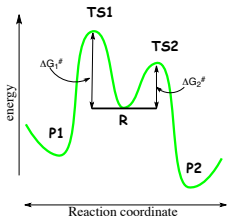


Figure 1

many processes where this is not the case. On the one hand, there is the kinetic control in which the reactant(s) face(s) a number of competitive reaction paths and the one with the lowest standard free energy of activation will be favored. (see Fig. 1). On the other hand it has also been observed that for many reactions (and the number of cases is growing),[1] kinetic selectively does not reduce to a simple choice between paths with different barrier heights and there are other factors, dynamical factors, coming into play to determine the relative abundance of the final products.

To described each of the "controlling factors" mentioned above, there exits some methodologies more suitable than others. Thus, for thermodynamically controlled processes the adequate electronic structure method is enough to discriminate between the different possible products and describe the factors favoring one product over the other(s). In this case, mechanism are irrelevant to the product ratio. For instance, this was the case for the properties studied in the previous section of this manuscript (Part I). For kinetically controlled reactions statistical theories such as transition state theory (TST) or Rice-Ramsperger-Kassel-Marcus (RRKM) theory are widely used in order to predict and explain the experimental product ratio. One of the main assumptions of these theories is that the energy that the molecule have to react (it is not important how the molecule did adquire it) is completely distributed within the internal modes of the molecule. When this assumption does not hold and the way of exciting the molecule, in that where the energy is initially located, is key to determine the reaction course and the mechanims involved, reactivity can only be understood by means of dynamical simulations.

Throughtout Part II such a case is presented, i. e. static methods fail to corretly, or fully, describe the reactivity of the molecules considered. Statistical approaches are able to partially account for that reactivity, but again, there are some loose ends. Dynamical simulation answer the questions that were not resolved with the previous approaches. However, if this method was to be used alone, it would have also been imposible to completely explain the reactivity on the process of interest. Thus, it is the combination of the three approaches that allows to correctly explain the phenomenon under survey.

PART I

1 Amine-boranes

Perhaps, the most prototypical of Lewis acid–base reactions is those of amines with boranes.[3,4] Gay-Lussac reported the successful preparation of H_3N–BF_3 in 1809; the history of N–B donor–acceptor complexes is thus nearly 200 years old.[5] Since 2000, these species have attracted increasing attention as reagents, hydrogen storage materials, polymer precursors, and also with respect to their coordination chemistry.[6] Primary amine-boranes can undergo dehydrogenation of dehydrocoupling reactions,[7]

Amine–adducts can be considered as textbook examples of Lewis acid–Lewis base adducts. Compounds containing dative bonds between boron and nitrogen have been known since the early 19th century. The first compound of this type, ammonia-trifluoroborane, H_3N–BF_3 was prepared in 1809 by Gay-Lussac[5] and also represented the first coordination compound of any type. Since this initial report, research on this adduct[8] and a plethora of analogs has been reported with important pioneering contributions by Wiberg and Stock.[9]

After discovery of ammonia-trifluoroborane a long period of time, until 1937, past before the first amine–borane containing only hydride of substituents on boron, $Me_3N{\cdot}BH_3$, was sinthesized.citeMe3-BH3 Since its synthesis by Burg and Schlesinger a rapid growth in the number of reported primary, secondary, and tertiary amine–borane adducts was initiated. In 2006, sterically hindered amino-boranes were shown to exhibit unexpected reactivity, including the heterolytic cleavage of various bonds, most notably that of dyhidrogen.[10] There are a number of technical applications of amine–borane adducts as stabilizants in polymer formulations, the bleaching of wood pulp, photographic applications, and as fuel additives.[11] Historically and to the present day, their most widespread applications are based on their reducing ability, either for uses in organic reactions or in electroless plating process, or

as easy to handle borane reagents for hydroborations. On the other hand, the use of amine–boranes as precursors to inorganic polymers and as interesting ligands with novel bondign modes is a very recent development.[6] The use of ammonia–borane as a potential protable hydrogen storage material has attracted a surge of interest as a result of the high hydrogen content (19.6 wt %).[12]

Based on the polarity of the N–B bond amine–boranes may have applications as liguid crystal lead strucutres.[13] Although not a direct application of amine–borane themshelves, the formation of a borane adduct of an existing amine can be used to protect the amine and often enables alternative chemistry to be performed. For instance, if the masking of an amine's nucleophilic properties is required, borane can serve as a protecting group. Furthermore, complexation of an amine with a borane may lead to a change in selectivity for further reactions with the amine.[14] Creo que hay mil millones mas de aplicaciones, que no voy a listar porque ya me estoy aburriendo. Creo que ha quedado claro que los amino-boranos son moleculas muy interesantes con trescientos mil millones de aplicaciones.

Novel coordination modes for amine–borane adducts are of fundamental interest for the understanding of catalytic dehydrogentaion, dehydrocoupling, or cross-couplig reactions, thereby enhancing the potential for improved system design. A fundamental understanding of the intrinsic properties of such adducs allows to improved the desygn of new system with the required properties for the desired application. The potential to gain further insight into the bonding in main-group complexes provides further motivation for the investigation of this intersting class of compounds.

Compounds formed by the association of amines and boranes have been known since the nineteenth century. [ref1,2] During the last decade such species have gained more attention as possible hydrogen storage materials and reagents in organic chemistry. [ref3,4] Recent research has revealed several interesting properties most notably that of the dihydrogen bond [ref5-7] and related "molecular tweezer" effect. [ref8-11] So far, the vast majority of amine-borane publications have been focusing on the catalyzed removal of hydrogen. Also, the effects of halogen and methyl substitutions on complexation energies and proton affinities have been widely studied. [ref12-17] Bulkier substituents and especially their effect on deprotonation reactions have received much less attention. Studies of the gas-phase proton transfer reactions give information about intrinsic properties and are therefore closely reltaed to possible applications.

There have been quite a few works concerning acidities of N-substituted amine-boranes. More than ten years ago Ren et al. investigated some borane complexes which included trimethylamine-borane and triethylamine-borane. [ref18] The experiments revealed that the acidity of trimethylamine-borane is between that of water and naphthalene. An acidity value of 384.0 ± 2.0 kcal mol^{-1}(1605.1 ± 8.4 kJ mol^{-1}) was assigned to the complex. They also estimated that the coordination of Me_3N by BH_3 leads to an α-CH acidity incrase of about 18 kcal mol^{-1}(75.2 kJ mol^{-1}). Experimental acidity for triethylamine-borane could not be determined due to an elimination reaction resulting in the removal of an ethyl group.

Very recently

Our aim is to show that intrinsic acidity is an important characteristic of amine-borane complexes. (Me imagino que con todos los ejemplos que ha puesto antes y que yo no he copiado). Herein, we show, using a combined experimental and theoretical study, that typical conventional bases such as aniline become acidic—with acidities as high as phosphoric acid— when they form complexes with borane, whereas the acidity of other nitrogen bases, such as dimethylamine, aziridine, and cyclopropylamine, becomes as high as that of formic, acetic, and propionic acids.

Let us start by making a brief comment on the methodology. The gas-phase acidity, deprotonation enthalpy, and deprotonation entropy of the amine–borane complexes have been experimentally determined by means of the "extended Cooks kinetic method " (EKM).[15–23] More details about the EKM method are given in the methodology section. Concerning the computational approach employed, we need to use a reliable model in order to analyze the electronic changes undergone by the Lewis acid and base when they interact. The first requirement to ensure the reliability of the model is a good agreement between the measured and calculated gas-phase acidities. This will provide an indirect evidence that the structures used in the calculations, for both the neutral and the deprotonated systems, are the same as those being probed experimentally. For this reason we used a high-level *ab initio* approach, that is, the one based on the G4 theory, which has been shown to provide very accurate values for the enthalpies of various reactions. [ref55] The G4 methodology may be too intensive when investigating very large systems. Thus, we have also assess the performance of a "cheaper" model, based on the use of the B3LYP[refB3LYP] density functional. This approach has been shown to perform very well for the calculation of the

intrinsic acidities of phosphine-borane complexes. [refmarcela] Within this model, the geometries are optimized using a 6-31+G(d,p) expansion and the final energies are obtained in sigle-point calculations using the aforementioned optimized geometries and a 6-311++G(3df,2p) basis set. All these calculations have been carried out with the Gaussian09 suite of programs. [refG09]

Since one of the main part in the bonding between amines and borane is the dative bond formed upon the transfer of electron density from the lone pair of the nitrogen atom of the amine to the empty $2p$ orbital of BH_3, the NBO approach, as implemented using the NBO-5G suite of programs,[24] is particularly well suited to describe these interactions. Whithin the NBO framework it is also possible to calculate the Wiberg bond order.[25] A complementary description of the bonding in amine-borane complexes can be obtained by means of the atoms in molecules[26] and the electron localization function (ELF)[27,28] theories. The AIM and ELF calculations were performed with the AimAll[29] and the TopMod[30] packages, respectively.

The calculated and measured gas-phase acidities of the free amines and the correponding amine-borane complexes, measured as the Gibbs free energy associated with the reaction 1 are summarized in Table 1. In all the cases, it is assumed that the proton is lost from the NH_2 group.

$$AH \rightarrow A^- + H^+ \tag{1}$$

For the sake of completeness, this table also includes the gas-phase acidities of the free amines. Only for some of the amines considered here are the experimental gas-phase acidities known. Unfortunately, the EKM method used in this work is not well suited to measure the acidity of very weak acids, such as the amines. Nevertheless, it is worth noting that for those cases in which the experimental acidity is known, the agreeement with our G4-based calculated values is excellent, and therefore, our estimates for the unknown gas-phase acidities should be accurate. This agreement between experimental and calculated values is also excellent for the gas-phase acidities of amine-borane complexes. The agreement is somewhat worse when the B3LYP values are used, values that are, in general, slightly lower than the experimental values. Nevertheless, there is a reasonably good linear correlation between the B3LYP/6-311+G(3df,2p) gas-phase acidities and the G4-based calculated values (Fig. 2), and therefore, in the case where the

Table 1: Experimental and G4-based calculated gas-phase acidities, ΔG^0_{acid} [kJ mol^{-1}], for several amines and the corresponding amine-borane complexes; $\Delta\Delta G^0_{acid}$ [kJ mol^{-1}] is the increase in acidity on going from the free amine to the amine-borane complex.

	Free amine		Amine-borane complex		
	ΔG^0_{acid} [kJ mol^{-1}]		ΔG^0_{acid} [kJ mol^{-1}]		$\Delta\Delta G^0_{acid}$ [kJ mol^{-1}]
Amine	exptl.[a]	calc.	exptl.	calc.[b]	calc.
ammonia	1656.8±1.6	1657.2	—	1462.1(1456.7)	195.1
methylamine	1651±11.0	1656.1	1461.0±9.2	1462.4(1455.2)	193.7
dimethylamine	1623±8.8	1621.8	1457.9±9.2	1453.7(1444.4)	168.1
allylamine	—	1616.5	1443.7±8.8	1444.2(1437.4)	172.3
cyclopropylamine	—	1618.3	1440.5±9.2	1447.3(1442.1)	171.0
benzylamine	—	1588.9	1436.7±8.9	1438.1(1435.5)	150.8
aziridine	—	1603.3	1443.4±8.9	1435.5(1432.2)	167.8
propargylamine	—	1608.8	1435.1±8.9	1431.1(1425.4)	177.7
trifluoroethylamine	—	1579.4	1405.0±9.4	1400.5(1393.5)	178.9
aniline	1502±8.4	1506.7	1365.7±9.4	1360.3(1353.6)	146.4

[a]Values taken from ref. [ref67]. [b] Values within parentheses were obtained at the B3LYP/6-311+G(3df,2p)//B3LYP/6-31+G(d,p) level.

gas-phase acidities of larger amine-borane complexes needs to be estimated, the theoretical model employed here can be used as a good alternative to G4 theory, which may be prohibitively expensive.

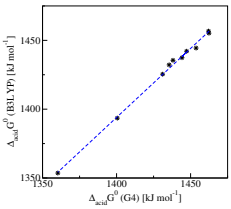


Figure 2

Active center for deprotonation. As previously mentioned, to compute the acidity of the amine-borane complexes it is assumed that deprotonation takes place in the amingo group of the complex. The anion so produced is in all cases, except for $PhCH_2NH_2{\cdot}BH_3$and $CF_3CH_2NH_2{\cdot}BH_3$, the most stable one. For both $PhCH_2NH_2{\cdot}BH_3$and $CF_3CH_2NH_2{\cdot}BH_3$, the most stable anions correspond to structures in which the proton is lost from the borane moiety (see Fig. 3 on the margin). The increased stability of these two structures, which can be viewed as the interaction between a NH_2BH_2 group and either the $C_6H_5CH_2^-$ or the $CF_3CH_2^-$ anions, respectively, just reflects the high

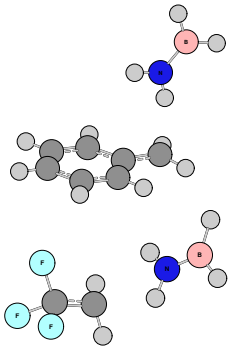


Figure 3

stability of both the neutral NH_2BH_2 group and the accompanying anions. Accordingly, these boron-deprotonated structures are predicted to be 22 and 15 kJ mol^{-1}more stable than the corresponding amine-deprotonated species, respectively. Nevertheless, the good agreement between the calculated and experimental values in Table 1 for these two amine-borane complexes seems to indicate that under experimental conditions only the amine-deprotonated species is formed.

To explain this apparent dichotomy we investigated in detail and compared BH_3 and NH_2 deprotonation by using $PhCH_2NH_2{\cdot}BH_3$ as a suitable example. As illustrated in Fig. 51, the most stable anion (structure **C**) is the result of the dissociation of the borane-depotonated species **B**, which involves a barrier (transition state, **TS_BC**) of 100 kJ mol^{-1}. However, direct deprotonation of the BH_3group of $PhCH_2NH_2{\cdot}BH_3$to yield structure **B** is much less favorable (by 300 kJ mol^{-1}) than the direct deprotonation of the amino group to yield anion **A**; this was found to be the case for all other amine-borane complexes investigated herein. It is also worth noting that the transfer of a proton from the BH_3group in anion **A** to the N atom, is accompanied by cleavage of the C-N bond. The consequence is that the transtition state associated with this proton transfer, namely **TS_AC**, directly connects anions **A** and **C**, through a barrier of 228 kJ mol^{-1}. Hence, in spite of its increased stability, form **C** can only be reached through a very energetically demanding processes, from either structures **A** or **B**.

In addition, an examination of the molecular electrostatic potential of both benzylamine and $CF_3CH_2NH_2$ shows that for both molecules the positive potential areas (blue) are those close to the amino group (Fig. 5). This means that the association of the molecule with the reference anion takes place at the amino group and never at the BH_3group; the interaction of the latter with the attacking anion would be highly repulsive because of the hydride character of the BH_3hydrogen atoms. These data are consistent with the much less favorable deprotonation of the BH_3group. Hence, under normal experimental conditions, the deprotonation of the amino group will be always favored and, as indicated above, its conversion into the more stable structure **C** would not take place because the transformation would involve a very high activation barrier.

Acidity enhancement. The first conspicuous fact is the significant increase of the acidity on going from the free amine to the corresponding

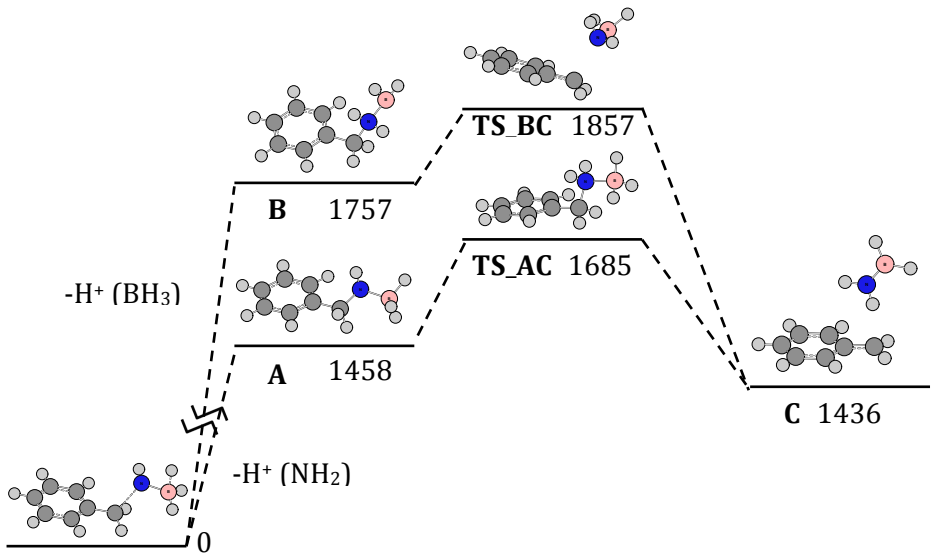


Figure 4: Energy profile of the NH_2 and BH_3deprotonation processes of benzylamine·BH_3. All values in kJ mol^{-1}.

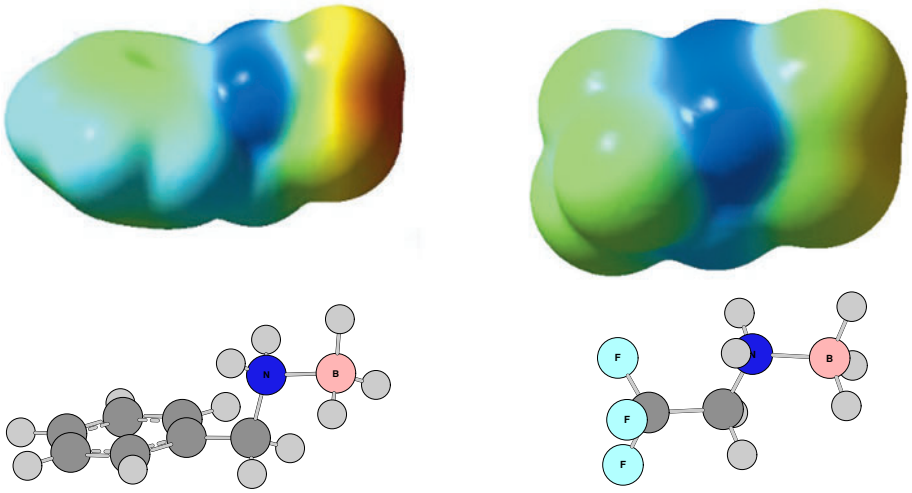


Figure 5: Molecular electrostatic potential of $PhCH_2NH_2$·BH_3(left) and $CF_3CH_2NH_2$·BH_3(right). Blue areas correspond to positive values of the potential, whereas red areas correspond to negarive values of the potential.

amine-borane complex (Table 1). The acidity enhancement is around 180 kJ $mol^{-1}$and depends on the nature of the group attached to the nitrogen atom. For instance, whereas deprotonation of ammonia and methylamine

borane leads to similar values of acidity enhancement ($\Delta\Delta^0_{acid} \approx 190$ kJ mol^{-1}), deprotonation of dimethylamine borane leads to an acidity enhancement about 25 kJ mol^{-1}lower. Also, smaller increases in acidity are observed for aniline and benzylamine complexes. Nevertheless, it is worth stressing that the increase in acidity that occurs upon coordination of the amines with borane is very large. Aniline borane has practically the same gas-phase acidity as phosphoric acid (1351 ±21 kJ mol^{-1}) [ref67] and most of the amine-borane complexes studied herein have gas-phase acidities similar to typical carbosylic acids, such as formic, ethanoic, and propionic acid. [ref67] Furthermore, the increase in acidity measured and calculated herein for amine-borane complexes is significantly larger than that measured and calculated for the phosphine-borane analogues. [ref11] For instance, whereas the increase in acidity observed for phenylphosphine and methylphosphine upon BH_3association is 78 and 123 kJ mol^{-1}, respectively, the increase in acidities for the amine-borane analogues are almost twice these values, 149 and 191 kJ mol^{-1}, respectively. These results prompted us to study whether the same phenomenon occurs when the Lewis acid is other than BH_3.

Conclusions. Before starting the following section let us briefly summarize the most important findings of this part. From our combined experimental and theoretical survey we conclude that the complexation of different amines with BH_3 leads to new compounds (amine-borane complexes), which exhibit a much larger gas-phase acidity. The acidity enhancement depends on the R substituent attached to the NH_2, going from 146 to 195 kJ mol^{-1}. The unexpected consequence is that typical nitrogen bases such as aniline, lead to amine-borane complexes which, in the gas-phase, are as strong an acid as phosphoric, oxalic, or salicylic acid, and stronger than many carboxylic acids, such as formic, acetic, and propanoic acids. The agreement between experimental and the G4-based calculated values is excellent. Also there is a good agreement between the G4 values and the values computed using the DFT approach.

2 Group 13 Lewis acids: BH_3, AlH_3 and GaH_3

The aim of this section is to investigate what would be the effect of replacing borane, BH_3, by alane (AlH_3) or gallane (GaH_3) in the amine adducts previously considered and phosphine analogues. Volatile Lewis base adducts of alane, as well as gallane, have an interesting chemistry and play an im-

portant role in a wide range of chemical process.[31–33] Particularly interesting are their applications on areas as material science . Aluminum hydride derivatives (and borane derivatives) can be used as precursors of AlN/BN materials.[34] Since their adducts have a volatile behavior, alane and gallane are used in chemical vapor deposition(CVD) technology.[35,36] Different complexes of AlH_3 with nitrogen bases have been proposed as potential hydrogen storage systems[37] as well as AlH_3PH_3.[38] Gallium hydride derivatives have been used as molecular precursors to thin films of Ga_2O.[39] Both gallium and gallium suboxide, Ga_2O, are used to make III-V semiconductors. Lewis base adducts of alano and gallano are also known to have important applications in organic synthesis as selective reducing agents[40,41] and in the preparation of transition metal aluminum and gallium hydrides.[42–44] Phosphine-alanes can act as catalysts in many chemical processes.[45] There is consequently a lot of interest in how aluminum and gallium molecules interact with other species.

In order to analyze the effect of AlH_3 and GaH_3 on the intrinsic acidity of amines and phosphines we have selected, as suitable model systems, three different R–NH_2 amines and the phosphorus containing analogues, R–PH_2. The R substituents have been chosen to represent three different situations: an alkyl group (methyl), a saturated non-aromatic ring (*c*-C_3H_5), and an aromatic ring (phenyl). To the aforementioned three amines (phosphines) we have added NH_3 (PH_3) as a good reference system and due to the numerous potential applications these molecules present. Thus we will compute the gas-phase acidities of $RNH_2{\cdot}AlH_3$, $RPH_2{\cdot}AlH_3$, and $RPH_2{\cdot}GaH_3$ Lewis complexes. Then, the values obtained will be compared with those previously reported for the analogoues amine-boranes and phosphine-boranes. In this way, we will try to understand the behavior of group 13 trihydrides compounds as Lewis acids.

As shown in the previous section, the B3LYP model is a good alternative to the use of G4 when computing gas phase acids of this kind of Lewis adducts. Thus, we optimized the geometries and computed the thermodynamical data at the B3LYP/6-31+G(d,p) level of theory and the final energies were obtained in sigle-point calculations at 6-311++G(3df,2p) level. The intrinsic acidities thus calculated for the free amines and the corresponding adducts are listed in Table 2. Table 3 lists the corresponding values for the phosphorus derivatives. In all the cases, it is assumed that the proton is lost from the heteroatom. In order to make easier the comparison between the different adducts, gas-phase acidities for amino-boranes and phosphino-boranes were also included in Tables 2 and 3, respectively.

Table 2: B3LYP-calculated acidity ($\Delta_{acid}G^0$, kJ mol^{-1}) and acidity enhancement ($\Delta\Delta_{acid}G^0$, kJ mol^{-1}) for R–NH_2 amines and the corresponding R–NH_2·BH_3 amine–boranes and R–NH_2·AlH_3 amine–alanes.

	$\Delta_{acid}G^0$			$\Delta\Delta_{acid}G^0$	
R	RNH_2	RNH_2·BH_3	RNH_2·AlH_3	RNH_2·BH_3	RNH_2·AlH_3
H	1656.3(1656.8±1.6)[a]	1456.7	1398.3	199.6	258.0
CH_3	1646.5(1651±11.0)[a]	1455.2	1406.3	191.3	240.2
c-C_3H_5	1610.9	1442.1	1386.6	168.8	224.3
C_6H_5	1503.1(1502 ±8.4)[a]	1353.6	1321.7	149.5	181.4

[a]Experimental value taken from ref.[46]

Table 3: B3LYP-calculated acidity ($\Delta_{acid}G^0$, kJ mol^{-1}) and acidity enhancement ($\Delta\Delta_{acid}G^0$, kJ mol^{-1}) for R–PH_2 phosphines and the corresponding R–PH_2·AlH_3 phosphine–boranes, R–PH_2·BH_3 phosphine–alanes, and R–PH_2·GaH_3 phosphine–galanes.

	$\Delta_{acid}G^0$				$\Delta\Delta_{acid}G^0$		
R	RPH_2	RPH_2·BH_3	RPH_2·AlH_3	RPH_2·GaH_3	RPH_2·BH_3	RPH_2·AlH_3	RPH_2·Ga
H	1504.1(1509.7±2.1)[a]	1384.0	1325.5	1341.7	120.1	179.2	163.0
CH_3	1533.3(1530.0±2.5)[b]	1410.2[c]	1359.3	1373.3	123.1[c]	174.0	160.0
c-C_3H_5	1515.4(1510.0±3.0)[b]	1402.0[c]	1352.1	1372.8	113.4[c]	163.3	142.6
C_6H_5	1457.0(1457.3±0.8)[b]	1379.1[c]	1324.6	1339.5	77.9[c]	132.4	117.4

[a]Experimental value taken from ref.[47] [b]Experimental values taken from.[48]
[c]Calculated values taken from ref.[48]

As found for amino-borane complexes, the adducts condisdered in this section behave as heteroatom (N or P) acids in the gas phase. For example, for the CH_3PH_2·AlH_3 adduct, the C–H deprotonation and the Al–H deprotonation processes are 138 and 201 kJ mol^{-1}, respectively, less favorable than the P–H deprotonation. The corresponding C–H and Ga–H deprotonation of the CH_3PH_2·GaH_3 adduct are 128 and 153 kJ mol^{-1}less favorable.

For all the molecules considered, and as in the case of the amine-borane complexes, a significant increase of the acidity of the Lewis base is systemat-

ically observed when the Lewis adduct is formed, both with alane, AlH_3, and gallane, GaH_3 (see Tables 2 and 3). Quite unexpected, however, the acidity enhancement predicted for the amine-alanes is significantly larger than that measured for the amine-boranes, around 30%. Accordingly, aniline–alane is predicted to be ca. 30 kJmol^{-1}more acidic than aniline–borane. The consequence is that aniline–borane would become a stronger acid in the gas-phase than phosphoric acid (1351±21 kJ mol^{-1})![46] Also the acidity enhancement predicted for AlH_3 when forming complexes with phosphines is larger, about 45% and 30%, than that induced by BH_3 or GaH_3, respectively. This is particularly the case for $PhPH_2$ where the acidity enhancement triggered by alane is 70% larger than that induced by borane. As for the amine-borane complexes, the acidity enhancement depends on the nature of the R group attached to the heteroatom (N,P). For instance, the acidity enhancement is notably larger for R = CH_3 or *c*-propyl than for R = Ph. From Tables 2 and 3 we also concluded that the increase of the gas-phase acidity upon Lewis acid comlexation is always larger for amines than for the phosphine analogues. Hence, for a same R, the calculated acidity enhancement follows the trend $RNH_2{\cdot}AlH_3 > RNH_2{\cdot}BH_3 > RPH_2{\cdot}AlH_3 > RPH_2{\cdot}GaH_3 > RPH_2{\cdot}BH_3$, for all R considered. These results raise the following questions: why is the effect of AlH_3 or GaH_3 on the intrinsic acidities larger than that observed upon BH_3 complexation? Why does this effect depend significantly on the nature of the R substituent? Or, why does this effect is different for amines than for phosphines? In order to answer this questions it is first necessary to understand why the acidity of the Lewis base is enhanced when it associates with the Lewis acid Y.

Origin of the acidity enhancement To explain the origin of the acidity enhancement we will use the thermodynamic cycle shown in Fig. 6.

In this thermodynamic cycle, the values of ΔG_1^0 and ΔG_2^0 measure the stabilization undergone by the free amine (or phosphine) and its conjugate base upon Lewis acid, Y, complexation. Accordingly, ΔG_3^0 and ΔG_4^0 are the gas-phase acidities of the free amine (phosphine) and the corresponding Lewis adduct, respectively. Hence, the fact that the acidity of the Lewis adduct is larger than the acidity of the free Lewis base ($\Delta_r G_4^0 > \Delta_r G_3^0$) necessarily implies that the stabilization of the deprotonated species upon Lewis acid complexation is larger than that undergone by the corresponding neutral free amine (phosphine), $|\Delta_r G_2^0| > |\Delta_r G_1^0|$, see equation 2. The values for the stabilization upon complexation for the neutral and deprotonated forms

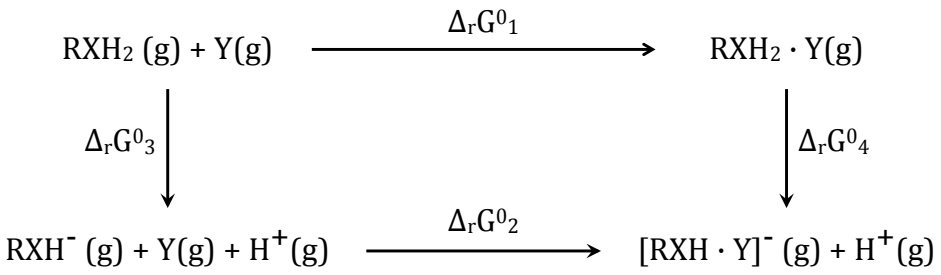


Figure 6: Thermodynamic cycle relating the intrinsic acidities of the isolated Lewis bases, RXH_2, X = N, P, ($\Delta_r G_3^0$) and the corresponding Lewis adducts, $RXH_2{\cdot}Y$, Y = BH_3, AlH_3, GaH_3, ($\Delta_r G_2^0$) with the stabilization undergone by the neutral ($\Delta_r G_1^0$) and deprotonated ($\Delta_r G_4^0$) forms of the free and complexed molecules.

of the free amine (phosphine) are listed in Table 4. The corresponding values for the remaining amino-borane molecules computed are shown in Table A.12 on the Appendix.

$$\begin{aligned} \Delta_r G_1^0 + \Delta_r G_4^0 &= \Delta_r G_3^0 + \Delta_r G_2^0 \Longrightarrow \\ \Longrightarrow \Delta_r G_2^0 - \Delta_r G_1^0 &= \Delta_r G_4^0 - \Delta_r G_3^0 = \Delta\Delta_{acid} G^0 \end{aligned} \quad (2)$$

Table 4: Stabilization free energies of neutral ($\Delta_r G_1^0$) and deprotonated ($\Delta_r G_2^0$) amines, R–NH_2, and phosphines, R–PH_2, upon BH_3, AlH_3 and GaH_3 complexation. All values are in kJ mol^{-1}.[a]

	$\Delta_r G_1^0$					$\Delta_r G_2^0$				
	R–NH_2		R–PH_2			R–NH_2		R–PH_2		
R	BH_3	AlH_3	BH_3	AlH_3	GaH_3	BH_3	AlH_3	BH_3	AlH_3	GaH_3
H	-77.9	-63.4	-49.4	-23.2	0.25	-273.1	-321.4	-169.6	-205.5	-162.8
CH_3	-96.3	-68.7	-55.7	-43.3	-15.2	-290.1	-308.9	-176.2	-217.6	-175.3
c-C_3H_5	-71.2	-56.8	-56.0	-42.6	-13.6	-251.3	-281.1	-163.2	-203.2	-156.2
C_6H_5	-57.7	-41.1	-50.8	-37.1	-7.2	-209.1	-222.5	-128.8	-170.1	-124.7
average	-75.8	-57.5	-53.0	-36.6	-8.9	-255.9	-283.5	-159.4	-199.1	-154.7

[a]Values for complexation with BH_3 taken from ref,[48] except for R = H which corresponds to the calculations of this work.

Indeed, values in Table 4 point out that the stabilization upon Y complex-

ation of the anionic species is significantly higher, one order of magnitude, than the stabilization undergone by the neutral amine (phosphine). This fact can be easily rationalized by taking into account that deprotonated amines (phosphines) are much better electron donors than their neutral counterparts. This is not surprising since the formation of the anion triggers a significant increase of the orbital energies, and in particular the energy of the highest occupied molecular orbital (HOMO), increasing its electron-donor capacity. This change on going from the neutral to the deprotonated species is nicely reflected in the NBO characteristics of the bond formed between the heteroatom, X (X = N,P), and the Lewis acid atom B, Al, or Ga. In Table 5 are listed the characteristics of the N–B bond, for some selected amine-borane complexes, in terms of atominc hybrids, Wiberg bond index, and electron density (ρ_b) at the N–B bond critical point (BCP). Similar tables for the rest of the molecules considered in this work can be found in the Appendix (Tables A.13, A.14, A.15, and A.16).

Table 5: Characteristics of the B–N bond in amine-boranes in terms of atomic hybrids, Wiberg bond index, and electron density (ρ_b(B–N), a.u.) at the corresponding BCP.

	Bond		Wiberg bond index		$\rho_b(N-B)$	
Amine	neutral	anion	neutral	anion	neutral	anion
Ammonia	82% **N**(35%s + 65% p)+ 18% **B**(16%s + 84% p)	75% **N**(40%s + 60%p) + 25% **B**(23%s + 77%p)	0.618	0.816	0.105	0.147
Methylamine	82% **N**(34%s + 66% p)+ 18% **B**(16%s + 84% p)	75% **N**(39%s + 61%p) + 25% **B**(23%s + 77%p)	0.612	0.785	0.112	0.151
Dimethylamine	82% **N**(32%s + 68% p)+ 18% **B**(16%s + 84% p)	76% **N**(37%s + 63%p) + 24% **B**(22%s + 78%p)	0.592	0.745	0.114	0.150
Aniline	LP(**N**)→ 2p (**B**)[a] 1079	77% **N**(43%s + 57%p) + 23% **B**(21%s + 79%p)	0.608	0.743	0.109	0.140

[a]The value reported corresponds to the second order interaction energy, in kJ mol^{-1}, between the N lone pair (LP(**N**)) and the empty $2p$ orbital of **B** atom.

From the data on Table 5 it can be observed that the donation of the lone pair of the nitrogen atom into the empty $2p$ orbital of the boron atom leads to a strongly polar chemical bond in which the contribution (82%) of the nitrogen-based hybrid orbitals to the bond is dominant. For the deprotonated species, the contribution of the boron-based hybrid orbitals to the

bond is significantly higher (from 18% to 24%) than in the neutral species. A concomitant increase of the *s* character of the hybrid orbitals (from both, N and B) participating in the bond is also observed. Consequently, the bond between the boron atom and the nitrogen atom in the deprotonated amine is stronger than in the neutral complex. This fact is also reflected in both the value of the Wiberg bond index which is about 27% larger in the anion, and the value of the electron density, ρ_b, at the corresponding bond critical point (see Table 5 and Ap.13 in the Appendix). Same trends are observed for amine-alanes and phosphine-alanes, with the Al contribution to the N(P)–Al bond going from 8% in the neutral amine-alane to 13% in the anion, and from 13% to 24% in the case of phosphine-alanes. Similar values are found for phosphine-gallanes. An increase of the *s* character of the orbitals participating in the bond is also observed for the deprotonated forms of these complexes (see Tables Ap.14, Ap.15, and Ap.16 in the Appendix).

The ELF plots (Fig. 7) are consistent with the previous analysis. These plots show how the lone pair that is created upon deprotonation of the amino group connects (partially delocalizes) with the disynaptic B–N basin. This effect becomes more apparent in the case of aniline, where the new nitrogen lone pair and the B–N disynaptic basin appear as an unique basin of population 3.64 electrons.

Differences between Lewis acids. Turning now to the differences between distinct Lewis acid, the greater acidity enhancement effect observed for alane and gallane with respect to borane complexes is in principle unexpected taking into account that BH_3 should behave as a better Lewis acid than AlH_3 or GaH_3. In fact, the dissociation energies of the complexes of AlH_3 with PF_3, PCl_3, PMe_3, and $P(CN)_3$ were reported to be smaller than for the analogous complexes in which AlH_3 is replaced by BH_3.[49] Indeed, this is also observed when looking at the stabilization of the neutral amines and phosphines under investigation here ($|\Delta_r G_1^0|(BH_3) > |\Delta_r G_1^0|(AlH_3) > |\Delta_r G_1^0|(GaH_3)$) (see Table 4). Why however, is the stabilization of the anion greater when it donates to AlH_3 than when it does so to BH_3 or GaH_3? The explanation is nicely visualized in the corresponding ELF plots (Fig. 8)

Comparing the ELF of $PH_3{\cdot}Y$ (Y = BH_3, AlH_3 and GaH_3) complexes and their corresponding phosphorus-deprotonated species we observe that, for the neutral complexes the differences are minimal. For the deprotonated species there is a clear delocalization of the lone pair created on the P atom upon its

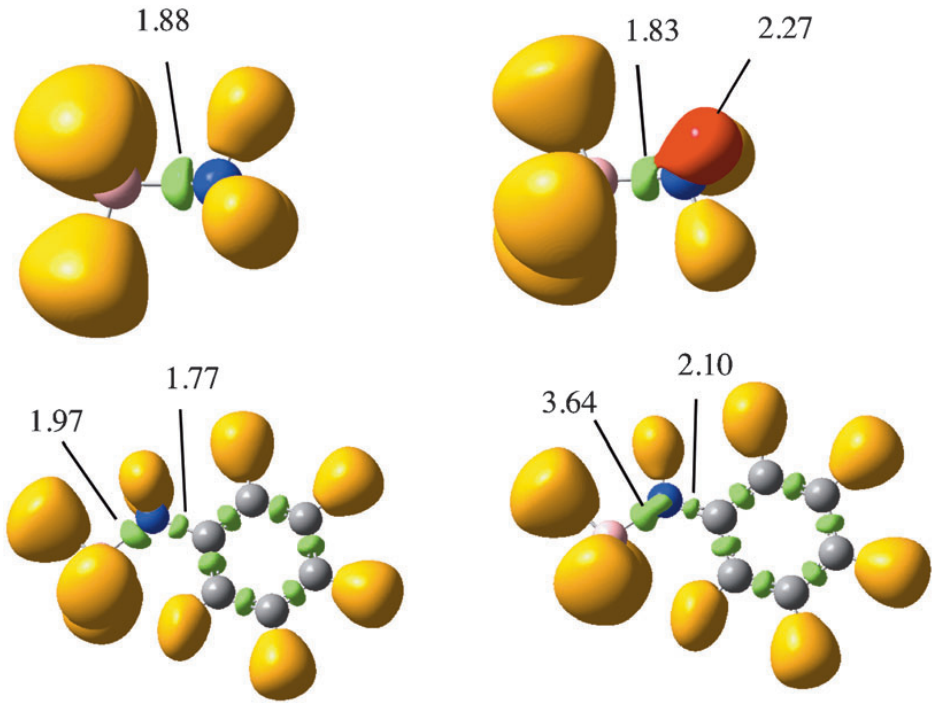


Figure 7: ELF (0.80) for the $NH_3{\cdot}BH_3$ and $C_6H_5NH_2{\cdot}BH_3$ complexes and their corresponding nitrogen-deprotonated species. Green lobes denote disynaptic basins involving two heavy atoms. Orange lobes are disynaptic basins in which H is one of the atoms involved. Red lobes correspond to the lone pairs. The populations shown are e^-.

deprotonation and the disynaptic P–Y (Y = B, Al, Ga) basin. This electron delocalization results in an increase electron flux to the bonding region. It is here where the differences appear. The electron delocalization, likely due to a lower electronegativity of the Al atom, is larger in the case of alane than for borane and gallane which is reflected in an larger electron population of the P–Al dysinaptic basin with respect to the population of the P–B and P–Ga disynaptic basins (1.93 e^- *vs.* 1.83 and 1.89 e^-). This fact is even more dramatic in the cases of the amines, where the population of the N–B disynaptic basin is 1.81 compared to 2.22 in the N–Al one. This fact would explain the enhanced stability of the $[XH_3{\cdot}AlH_3]^-$ (X = N, P) complexes with respect to that of $[XH_3{\cdot}YH_3]^-$ (X = N, P; Y = B, Ga). Therefore, the larger acidity observed for alane complexes can be attributed to a larger stabilization of the anionic complex compared to the borane and gallane complexes. It is known that the chemistry of gallane is more like borane than alane which

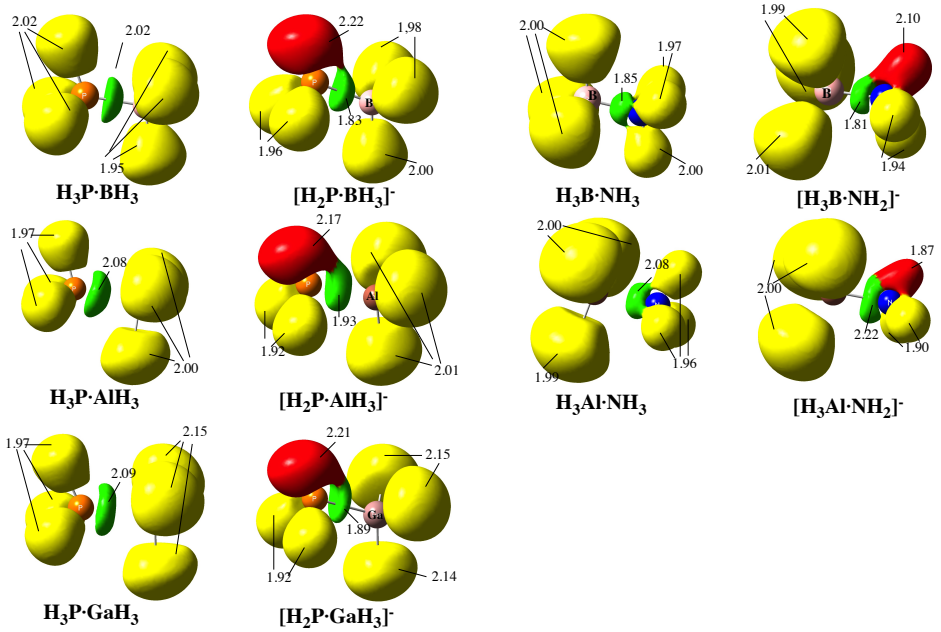


Figure 8: ELF (0.80) for the $XH_3{\cdot}Y$ (X= N, P; Y = BH_3, AlH_3 and GaH_3) complexes and their corresponding X-deprotonated (X= N, P= species. Green lobes denote disynaptic basins involving two heavy atoms. Yellow lobes are disynaptic basins in which H is one of the atoms involved. Red lobes correspond to the lone pairs. The populations shown are e^-.

is related to the periodic anomaly of the group 13 elements. Notably the increase in electronegativity of gallium over aluminium (1.8 Ga, 1.5 Al cf. 2.0 B, Allred Rochow) and the associated lower polarization influence of gallium.[33] Indeed, the stabilization undergone by the deprotonated phosphine upon BH_3 and GaH_3 complexation is almost the same. The average of the $\Delta_r G^0_2$ values computed is -159 and -155 kJ $mol^{-1}$for borane and gallane complexes, respectively. Also the values for the electron population of the P lone pair and P–B/Ga disynaptic basin in the deprotonated complexes are almost the same (see Fig. 8). The difference in electronegativity has already been pointed out in previous *ab initio* studies as the main factor distinguishing aluminum from its neighboring group 13 elements.[50]

Differences between R substituents. Let us compare the effect of the nature of the R substituent on the acidity enhancement. It can be seen that the acidity enhancement is sistematically lower for $PhXH_2$ (X = N, P) than

for the other molecules under study (Tables 1, 2 and 3). The reason is related to the aromatic character of the system. As shown in Table 4, the free aniline (phenilphosphine) is a poorer electron donor than the other amines (phosphines), as reflected in the lower ΔG_1^0 value. The reason is that the conjugation of the lone pair on the nitrogen (phosphorus) atom with the aromatic system. This is consistent with the fact that NBO analysis (see Tables 5 and Ap.13, Ap.14, Ap.15, Ap.16 in the Appendix) does not locate a B-N bond (with the default indexes of the NBO 5.0 program) but instead locates a dative bond between the lone pair of the heteroatom (N or P) and the empty p orbital on the B, Al or Ga atom. The anion is also significantly stabilized through the conjugation of the X (X = N, P) lone pair created in the deprotonation process with the aromatic system. Complexation of the anion with the Lewis acid, BH_3, AlH_3 or GaH_3, necessarily competes with this delocalization. Since a significant amount of the lone-pair charge is transferred to the Lewis acid molecule upon complexation, less is available for the conjugation with the aromatic ring. This is mirrored in a shortening of the C–X bond, as well as a significant increase of both the electron density at the corresponding BCP and the electron population of the C–X disynaptic basin (see Fig. 9). Still, the deprotonated amine (phosphine) is a better

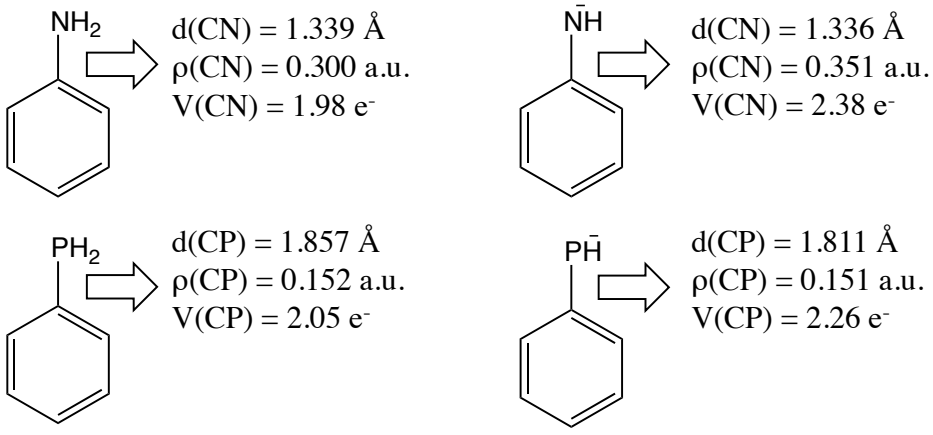


Figure 9: C–X (X= N, P) bond characteristics for Ph–XH_2 and their deprotonated couterparts.

electron donor toward the Lewis acid than the neutral species, through the second lone-pair created. However, the donor capacity of the second lone pair is smaller when R = Ph than in the other examples because the aforementioned conjugation decreases the intrinsic basicity of the XH_2 group.

The origin of the differences between methylamine and dimethylamine is more subtle. As shown in Table Ap.12 the stabilization of the neutral form (ΔG_1^0) is larger for dimethylamine than for methylamine as expected from the increase in the number of methyl substituents. However, the values for the stabilization of the anion ΔG_2^0 show the reverse order. This happens because the inductive effect of the methyl group makes dimethylamine a better Lewis base than methylamine. On the other hand, the extra methyl group significantly contributes to the dispersion of the excess of negative charge in the anion, enhancing its stability and attenuating the electron-donor ability of the nitrogen atom. The sligthly decrease of the acidity enhancement upon methyl substitution is also observed on going from $NH_3{\cdot}AlH_3$ to $CH_3HN{\cdot}AlHH_3$ and have been reported for methylamine/ammonia and methanol/water forming complexes with BeH_2 Lewis acid.[51]

Lewis base active site. As previously noted, the acidity enhancement effect is smaller when the active site of the base is a phosphorus. Also, the stabilization energies of both the neutral and deprotonated forms are smaller for the phosphorus bases. If we compare the NBO characteristics of the H_3N–BH_3 with those of the H_3P–BH_3 bond, we find that the N–B bond has a very large weight of the N atomic orbitals (82%), whereas the weight of the P atoic orbitals in the P–B bond is quite smaller (66%), reflecting the smaller electronegativity of the P atom. The same is observed for the alane complexes (see Tables Ap.14 and Ap.15). The interaction between the empty p orbital of the boron atom and the lone pair of the nitrogen is stronger than that involving the lone pair of a phosphorus atom. This difference is presumably due to the large difference in the size of the orbitals participating in the interaction in the latter case. However, this can not be the reason for the larger stabilization of the neutral and anionic forms of amines than that for phosphines upon complexation with alane. For AlH_3 complexes another factor comes into play, the electrostatic character of the bond. In this case, it is the electrostatic component that makes N–Al interaction stronger than P–Al interaction. Indeed, whereas the natural charges of the N and Al atoms in $H_3N{\cdot}AlH_3$ are -1.18 and +0.95, respectively, those of the P and Al atoms in the analogous adduct are +0.10 and +0.78, respectively. Figure 10 shows the natural charges for these two complexes. The coloring code goes from red for the most negative to green for the most positve. Black color correspond to zero (or close) charge. For the sake of comparison the corresponding borane complexes have been also plotted in the figure. It is apparent from Fig. 10 that the N–Al bond has the greatest electrostatic character of the four bond

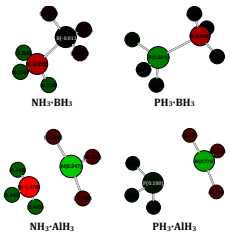


Figure 10

considered. These arguments explain why the stabilization energies (ΔG_1^0 and ΔG_2^0) are smaller for RPH_2 than for RNH_2, but do not explain why the acidity enhancement is also smaller for phosphine derivatives. Second-row atoms disperse the excess of negative charge much better than first-row atoms and thereby its electron-donor ability decreases. However, this effect is more pronounced in the anionic species. Thus, when going from amine to phosphines, the decrease in ΔG_2^0 (stabilization energy for the anion) is larger, in relative terms, than the decrease in ΔG_1^0 (stabilization energy for the neutral). Consequently, the difference between ΔG_1^0 and ΔG_2^0 is smaller in the phosphorus bases. This difference is the acidity enhancement.

In order to stress the significant acidity enhancement observed in the Lewis bases considered upon **Y** complexation let see the effect of the aforementioned acidity enhancement of the acid dissociation constant, k_a. An equilibrium constant is related to the standard Gibbs energy change for the reaction, $\Delta_r G^0$ through eq. 3

$$k_a = exp\left(-\frac{\Delta_r G^0}{RT}\right) \tag{3}$$

where R is the ideal gas constant and T stands for the absolute temperature (298 K). Because of the exponential relation between k_a and $\Delta_r G^0$ even small changes in the latter produce a large change in k_a. Thus, the acidity enhancements observed in the molecules studied suppose changes in the acidity constant ranging from 21 to 45 orders of magnitude!

Conclusions. The acidity enhancement effect of BH_3 observed in the aminoborane complexes is also found for other Lewis acids of the group 13, AlH_3 and GaH_3. This acidity increase is huge, going from 78 kJ $mol^{-1}$for the $PhPH_2{\cdot}BH_3$ adduct to 200 kJ $mol^{-1}$for $NH_3{\cdot}$ AlH_3. In terms of the acidity constant it supposes a change of 21 to 45 orders of magnitude!

The origin of the acidity enhancement is the larger stabilization of the anion than that undergone by the neutral basis upon Lewis acid complexation. The analysis of the electron density rearrangements on both the Lewis base and the Lewis acid show that the deprotonated base is a much better electron donor than the neutral one. Furthermore, the contribution of the Lewis acid to the dispersion of the excees electron denstiy of the anion is another factor contributing to the increased stability of the anions and therefore, to the increased acidity of the Lewis adducts with respect to the isolated bases.

This acidity enhancement depends on several factors: i) the Lewis acid considered, being larger for AlH_3 than for BH_3 and GaH_3. ii) the nature of the R substituents on the Lewis base, being smaller for R = Ph because of the conjugaton of the X lone pair with the aromatic system. iii) the Lewis base active site, being larger for amines than for phosphine derivatives. Hence, the calculated acidity enhancement follows this order: $RNH_2{\cdot}AlH_3 > RNH_2{\cdot}BH_3 > RPH_2{\cdot}AlH_3 > RPH_2{\cdot}GaH_3 > RPH_2{\cdot}BH_3$, for all R considered. All the molecules considered behave as N or P acids.

3 Group 15 Lewis bases: unsaturade derivatives

Up to now we have seen that when typical Lewis bases as amines and phosphines form complexes with group 13 Lewis acids, BH_3, AlH_3 or GaH_3, their intrinsic acidity changes dramatically, in most cases thirty orders of magnitude in terms of the equilibrium acidity constant. These acidity enhancements have been shown to be even larger when the Lewis acid is a beryllium derivative,[51] an electron deficient system, that as borane or alane, behaves as a very strong Lewis acid. The hard Be^{2+} cation (31pm) is very similar to Al^{3+} (50pm) in its charge to size ratio (0.0645 compared to 0.600, respectively). It has typically been compared to aluminium and is even paired with aluminum in recent inorganic chemistry texts.[52] Beryllium was originally discoverd by the French chemist Vauquelin in 1797 as a component of the mineral beryl and the gem emerald, and metallic beryllium was later isolated in 1828 independently by Bussy and Wohler.[53] Interest in the use of beryllium began in 1920 when it was demonstrated that a 2% addition of beryllium to copper produced an alloy six times stronger than the original material.[54] Since then, and clearly due to its unique properties that are a great asset in nuclear, aerospace and electronic industries, since beryllium has found extensive use in modern industry. A wide variety of applications have been developed, ranging from aircraft landing gear bushings to undersea telephone cable housings; and from oil field drilling equipement to golf clubs [ref2,3]. Therefore, there is much interest in further develope our understanding of the coordination chemistry of beryllium. However, the chemistry of beryllium compounds is much less explored than that of the other elements of the first-row of the Periodic System. Some coordination complexes of Be with nitrogen donor ligands have been reported, and few beryllim compounds of second-row elements are known. [refCottom] This comes from the

fact that beryllium compounds are quite toxic which makes experimental studies difficult. Indeed, beryllium is the most toxic non-radioactive element in the Periodic Table [ref1]. Skin poisoning requires surgical removal and, in severe cases, amputation may be required. [ref10] Inhalation of beryllium-containing dusts produces chronic pulmonary granulomatosis (berylliosis) or nodules in the lung [ref de el de 2001:4,6,9,21] The condition develops slowly and it is often fatal. Beryllium is lethal at 1ppm of body weight. Due to the *exceedingly toxic* nature of beryllium compounds, predictive theoretical calculations are specially interesting regarding Be chemistry. Thus, in this section we have considered BH_3, AlH_3 and BeH_2 as Lewis acids.

With regard to the Lewis base moiety, there are some evidences that seem to indicate that the acidity enhancement effect also depends on the nature of the Lewis base active site, since in general it has been found that the acidity enhancement is larger in amine adducts than in phosphine adducts, regardless the Lewis acid considered. Hence, in this section, we perform a more thorough study with regard to the Lewis bases. We expand the scope considering as Lewis bases group 15 derivatives, in particular molecules in which the heteroatom is directly bonded to an unsaturated moiety. More specifically to a vinyl or ethynyl group. We have also included the series of the saturated ethyl derivatives as a suitable reference. In Fig. 11 are shown the structures of the Lewis bases to be considered.

The aim of this section is to investigate the acidity trends of the Lewis bases down the group. Then, we will analyze the effect that the association of these compounds, with BH_3, AlH_3, or BeH_2 have on their intrinsic acidities. At this point, we will explore if the acidity changes are directly related with the deformation of the base, the acid or both. Finally the possibility that this association may lead to a change in the nature of the group loosing the proton will also be analyzed. The methodology used will be the same as detailed in previous sections.

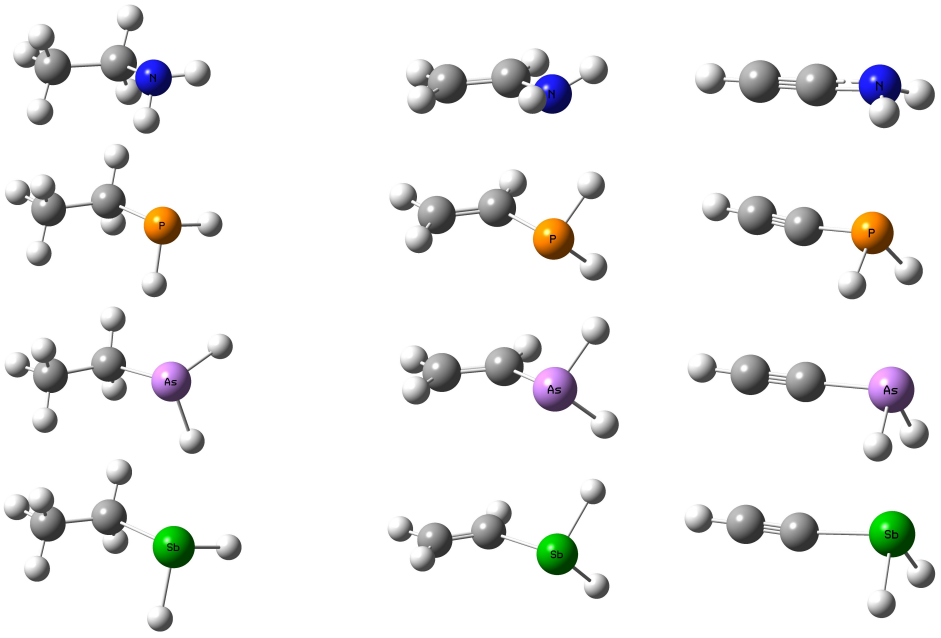


Figure 11: Structures for $H_3C–CH_2XH_2$, $H_2C{=}CHXH_2$, and $H_2C{\equiv}CHXH_2$ (X = N, P, As ,Sb) Lewis bases considered in this section.

4 Results

Acidity enhancement The calculated intrinsic acidities, measured as the Gibbs free energy associated with the reaction 1 are presented in Table A.17. In order to better visualize the trends going down the group these date are plotted in Fig12.

For the free compounds we have considered deprotonation at all possible acidic sites: X, C_α, and C_β. In all cases the systems behave as heteroatom acids since independently of the nature of R or X, the R-XH^- anion was found to be always the most stable one. This is in agreement with what has been found before for several amines and phosphines[48,55] and in the previous sections, as well as for vinylarsine,[55,56] vinylstibine[57] and ethynylarsine.

Again we want to note the very good agreement between our calculated acidities and the experimental values, whenever available (see Table Ap.17). Also the estimates of our DFT model are in excellent agreement with the G4 calculated values. It is also apparent that the acidity increases (lower values of $\Delta_{acid}G^0$) down the group for the three families of compounds, although this effect is attenuated on going from the saturated compounds to the vinyl

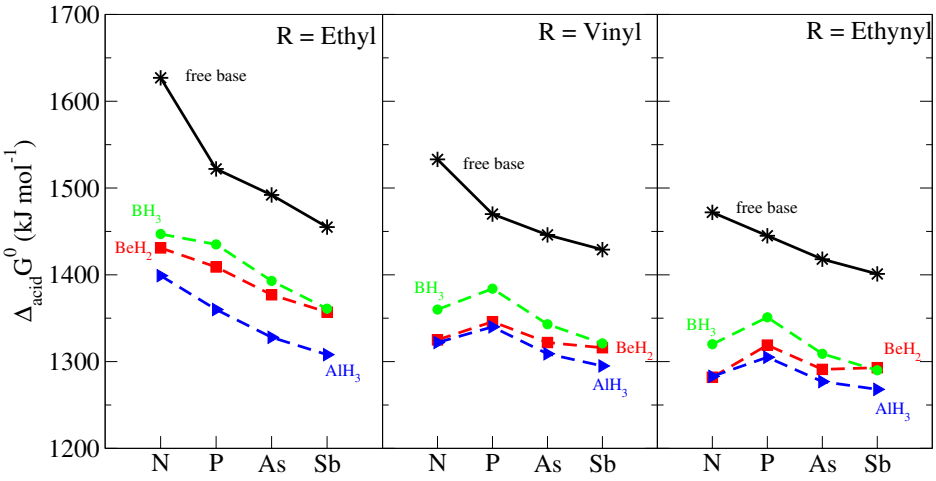


Figure 12: Calculated acidity ($\Delta_{acid}G^0$, kJ mol^{-1}) for R-XH_2 (R = Ethyl, vinyl, ethynyl; X = N, P, As, Sb) bases (solid lines) and the corresponding R-XH_2·BeH_2, R-XH_2·BH_3, and R-XH_2·AlH_3 complexes (dashed lines).

derivatives and further to the ethynyl ones. Indeed, while ethylstibine is predicted to be 172 kJ mol^{-1}more acidic than ethylamine, for the vinyl and the ethynyl analogues this gap is only 104 and 71 kJ mol^{-1}, respectively. As expected, the acidity increases as ethyl < vinyl < ethynyl, reflecting the larger electronegativity of the unsaturated groups wiht respect to the saturated one.

The complexation of the compounds under investigation whether it is with BeH_2, BH_3 or AlH_3 leads to a signigicant enhancement of their intrinsic acidities, similar to the ones reported in previous sections for other compounds. The largest calculated acidity enhancement, that can be seeing in the figures as the gap between the solid line and the dashed line for each Lewis acid, is found for aluminium complexes, followed by beryllium dihydride complexes and borane complexes. It is worth to note the similarity between BeH_2 and AlH_3 complexes of the unsaturated moieties. As aforementioned, some resemblances between the chemistries of Be and Al have already be observed.[3]

Again, we will use the thermodynamic cycle shown in Fig. 6 to rationalize these findings. With this purpouse the values of the stabilization energies of the neutral and deprotonated Lewis bases upon complexation, $\Delta_r G_1^0$ and $\Delta_r G_2^0$, respectively, have been listed in Table Ap.18. In Fig. 13 we show the corresponding graphs.

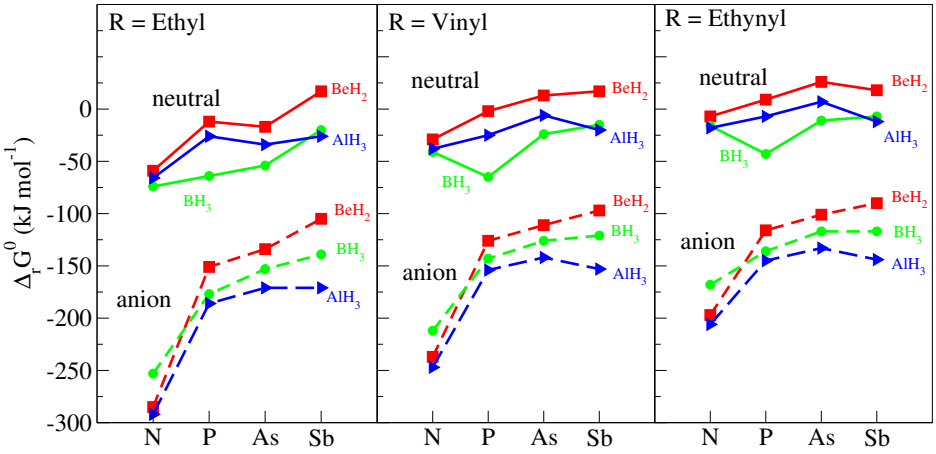


Figure 13: Stabilization free energies of neutral, $\Delta_r G_1^0$, (solid lines) and deprotonated species, $\Delta_r G_2^0$, (dashed lines) when R-XH_2 (R = Ethyl, vinyl, ethynyl; X = N, P, As, Sb) bases interact with BeH_2, BH_3, and AlH_3. The acidity enhancement, $\Delta\Delta_{acid}G^0$ corresponds to the gap between the solid and dashed line curves for the same Lewis acid. All values are in kJ mol^{-1}.

The values of $\Delta_r G_1^0$ and $\Delta_r G_2^0$ indicate that for the three kinds of complexes the acidity enhancement is due to a much larger stabilization of the anion than the neutral, when associated with any of the three Lewis acids. Actually, for the three BH_3 and AlH_3 and BeH_2 the strength of the X–Y (X = N, P, As, Sb; Y = B, Al, Be) bond dramatically increases on going from the neutral complex to the deprotonated one. This is well reflected in both, the values of the electron densities at the N–X BCPs and in the Wiberg bond orders. As shown in Fig. **??**, using the complexes with vinylamine as suitalbe examples, upon deprotonation of the complex, the electron density at the N–B, N–Al and N–Be BCPs increases by 0.039 0.025 and 0.034 a.u., respecitvely. Consistently, the N–B, N–Al and the N-Be Wiberg bond orders also increase from 0.526 to 0.572 for the BH_3 adduct, from 0.252 to 0.436 in the AlH_3 complex, and from 0.259 to 0.526 for the BeH_2 containing one.

It can be seen in Fig. 12, and values in the first three columns of Table Ap.**??** which give the magnitude of the acidity enhancement, that systematically, the effect is stronger when the Lewis acid is AlH_3. However, in general, the stabilization of the neutral base follows the order B > Al > Be, whereas the trend change for the anion: Al > B > Be (except for X = N), as can be seen in Fig. 13. Therefore, the larger acidity enhancement observed for alane complexes comes from a larger stabilization of the anion respect to the

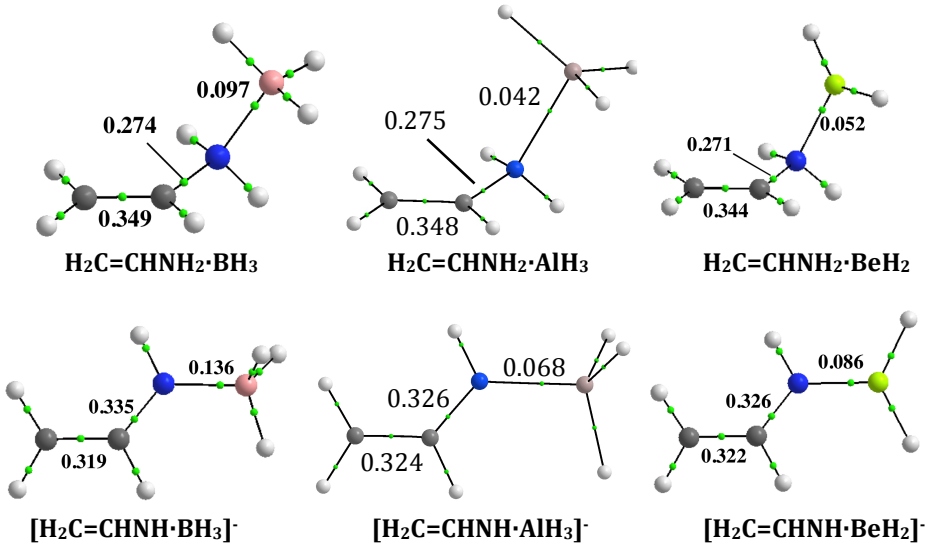


Figure 14: Molecular graphs of the BH_3, AlH_3 and BeH_2 complexes with vinylamine and their corresponding anionic deprotonated species. Green dots denote BCPs. Electron densities are in a.u.

neutral base compared to those of observed for borane. For BeH_2, even if in general the stabilization of both the neutral base and the deprotonated form is lower than for borane, the difference between $\Delta_r G_1^0$ (neutral) and $\Delta_r G_2^0$ (anion) is larger than for the latter. Hence, the acidity enhancement for BeH_2 complexes is larger than for BH_3 adducts. Indeed, from the values in Fig. 14, it can be seen that whereas the electron density at the N–Be BCP increases 65% upon deprotonation of the complex, that at the N–B BCP only increases by 40%.

There are other subtle differences between the three series of data. While for the BH_3 complexes the value of $\Delta_r G_1^0$ is always negative, i.e., the neutral compound is always stabilized by association with borane, this is not always the case upon association with BeH_2 or AlH_3. In principle, as shown in Table Ap.18, the formation of BeH_2 complexes for ethylstibine, vinylarsine, vinylstibine, ethynylphosphine, ethynylarsine and ethynylstibine are predicted to be endergonic processes, even though they are exothermic in terms of enthalpies. Since the reaction free-energies are rather small, we decided to verify whether these predictions could be an artifact of the DFT approach used, so for the P and As containing complexes the values of $\Delta_r G_1^0$ were re-evaluated at the G4 level. For the Sb containing compounds, for which

the G4 procedure is not available, the *ab initio* reference calculations were carreid out at the CCSD(T)/Def2-QZVP level of theory. These high-level *ab initio* values showed that, although the B3LYP method slightly underestimate the stability of the neutral beryllium complexes, the formation of the complexes of BeH_2 with ethynylarsine and ethynylstibine are indeed slightly endergonic (see Table A.18). This moved us to explore the relative stability of complexes in which BeH_2 interacts with the double or triple CC bond rather than with the heteroatom (see Fig. **??**).

This survey, carried out at the G4 level, showed that the global minimum for the interaction between both ethynylphophine or ethynylarsine and BeH_2 corresponds to a π-type complex, which was found to be 4 and 16 kJ mol^{-1}lower in free energy than the complex in which BeH_2 interacts with the P or the As atom, respectively. However, a similar survey for the complexes involving the analogous vinyl derivatives showed that always the complex in which BeH_2 is directly attached to the heteroatom, is more stable than the π-complex (29 kJ $mol^{-1}$for P and 13 kJ $mol^{-1}$for As).

It is worth noting that although for BH_3 containing complexes the value of $\Delta_r G_1^0$ for the ethynyl containing systems is smaller in absolute value than for the vinyl containing analogues, in all cases the formation of the complexes is predicted to be exergonic. Same applies for AlH_3 adducts, with the only exception of $HC{\equiv}C–AsH_2{\cdot}AlH_3$, with is predicted to be sligthly endergonic (7 kJ mol^{-1}).

Acidity trends Besides the acidity enhancement discussed above, the complexation also results in a change in the acidity trends. Whereas, as we have indicated before, the acidity of the free compounds increases down the group, the values of $\Delta_{acid}G^0$, for the BH_3, AlH_3 and the BeH_2 complexes, present a maximum (minimum acidity) for the vinyl and the ethynyl phosphine (dashed lines in Fig. 12).

Why the complexes of vinyl- and ethynyl-phosphine are less acidic than the corresponding amine complexes? Or, put in another way, why is the conjugate base of the amines more stable, in relative terms, than the conjugate base of phosphines? It is due to a significant stabilization of the anion through the conjugation of the N lone pair, created in the deprotonation process, with the π-system of the unsaturated vinyl/ethynyl substituent. This is nicely reflected in the weights of the resonant structures (see Fig.16) of the

$H_2C = CH - \overset{\ominus}{X}H \cdot YH_3$ (*a*)

$H_2\overset{\ominus}{C} - CH = XH \cdot YH_3$ (*b*)

Figure 16

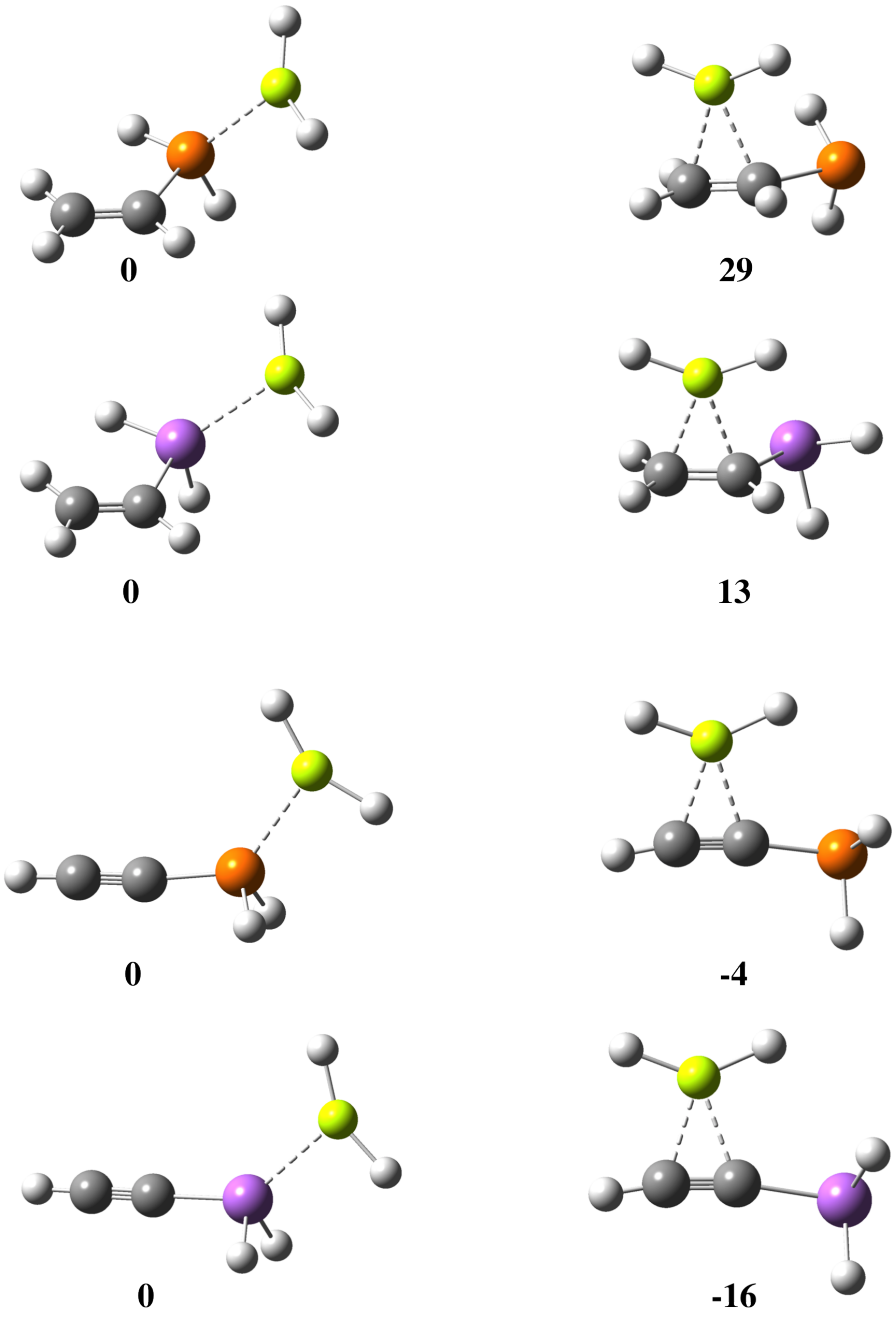


Figure 15: G4 relative stabilities (kJ mol^{-1}) for the more stable conformations of complexes of BeH_2 with vinyl- and ethynyl phosphine and arsine, showing that for ethynyl the π-type complex is the more stable.

neutral and anionic forms of the borane complexes with vinyl and ethynyl amines and phosphines (see Table 6).

Table 6: NRT weigths for *a*-type resonant structure of the vinyl–XH_2 and ethynyl–XH_2 (X = N, P, As, Sb) neutral and deprotonated BH_3 complexes.

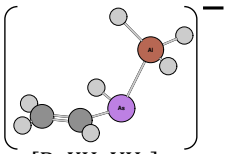
	Vinyl		Ethynyl	
	Neutral	Anion	Neutral	Anion
X = N	92%	60%	91%	60%
X = P	89%	86%	88%	87%
X = As	88%	87%	88%	88%
X = Sb	87%	90%	87%	91%

The results shows that whereas for the neutral complexes the resonant structure (*a*) is crearly predominant, there is not a strong conjugation, this is not the case for the anions. For both, vinyl and ethynyl derivatives, the weight of (*a*) structure for N containing bases decreases significantly, almost to half its value. However, this phenomenon is not observed for the heavier elements of group15, for which the weight of the resonant structure (*a*) remains almost invariable when the molecule is deprotonated. The greater acidity of As and Sb complexes than the N containing ones just reflects the decrease on the X–H bond strenghts going down the group.

The same phenomenon is observed in AlH_3, *i. e.*, it exits in the anions a strong conjugation of the amino group with the π-system of the unsaturated moiety, which is not observed for the phosphines. Several feature ratify this. The electron density at the C–N BCP increases dramatically when the anion is formed (0.275 $\rightarrow$ 0.326 a.u. (vinyl) and 0.308 $\rightarrow$ 0.348 a.u. (ethynyl)) and consequently the C–N bond length shrinks (0.08 and 0.06 Åfor vinyl and ethynyl). The double and triple character of the C–C bonds decreases, increasing the C–C distance (1.331 $\rightarrow$ 1.367 Åand 1.205 $\rightarrow$ 1.236 Å, respectively). It is worth to note that for ethynyl-amine a change in the HCC angle from linear to 159^o is also observed, showing the significant change in the hybridization of the C atoms. The behaviour is just the opposite for the P containg systems. The electron density at the C–P BCP decreases (0.161 $\rightarrow$ 0.151 a.u. and 0.161 $\rightarrow$ 0.148) and the C–P bond becomes longer (0.01 and 0.02 Åfor vinyl and ethynyl derivatives, respectively). The same kind of behaviour is found for the BeH_2 analogues complexes. Indeed, the

similarities between the values of the C–N and C–P BCP electron densities and bond distances are remarkably. All these results indicate that the N containing anions become much more stabilized in relative terms than the P containing ones, what should result in a much larger increase in the acidity for the amines than for the phosphines, explaining the appearance of the maxima in Fig. 12. Why the conjugation is favored in amines, is a well known mechanism, related to the efficienby of the overlap between the lone pairs of the heteroatom and the π-system when the heteroatom belongs to the first row.

Active center for deprotonation In our previous discussion it was implicitly assumed that for the Lewis complexes the proton lost would always come from the XH_2 group. This has been showed to be the case for different Lewis complexes as the ones previously exposed here and some others found in the literature (phosphine-boranes [ref 30] and complexes with BeH_2 derivatives. [ref34]) Also in the free compounds the proton is lost from the heteroatoms. Thus, in principle, this is a reasonable assumption also in this case. Neither the initial polarity of the bonds, for BeH_2, BH_3, and AlH_3 groups would favor the proton lost from any of these groups. For example, the hydrogens at nitrogen are aciditc in charcter because of the increased electronegativity of nitrogen, 3.04,[58] relative to hydrogen, 2.20.**??** Inversely, the hydrogens at boron are hydridic in nature because of the reduced electronegavitiy of boron, 2.04,[58] which leads to polarization of the bond toward hydrogen. Nevertheless, a thorough exploration of the potential energy surface reveals the existence of a stable anion in which the proton is lost at the Lewis acid site. Furthermore, this anion is in all the cases but for vinylamine derivatives, the most stable anion in terms of free energy (see Table 7). Furthermore, the formation of the anion at Be, B or Al site is accompanied by a significant structural rearrangement. The deprotonated Lewis group (BH_2, AlH_2, BeH) bridges between the C_{alpha} of the unsaturated moiety and the XH_2 group (X = N, P, As, Sb), so the anion is $[R–YH_2–XH_2]^-$ (insertion-anion) instead of $[R–XH–YH_3]^-$ (see figure 17). Similar structures to that shown in Fig 17 are found for all the complexes regardless the Lewis acid, Y, the heteroatom, X, or the nature of the R substituent.



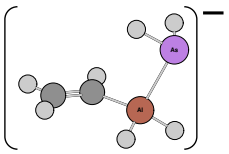


Figure 17

In light of the results showed in Table 7, the first thing that comes into mind is that the acidity enhancement reported in Table 17 is much larger for all the unsaturated systems investigated (with the only exception of vinylamine). Hence, in general the complexes of the ethynyl derivatives with

Table 7: Stabilities (kJ mol^{-1}) of the $[R–YH_2–XH_2]^-$ structures (insertion-anion) relative to those of the $[R–XH–YH_3]^-$ ones. R = vinyl-, ethynyl-; X = N, P, As, Sb; Y = Be, B, Al.

Y =	Be		B		Al	
R =	Vinyl	Ethynyl	Vinyl	Ethynyl	Vinyl	Ethynyl
X = N	+3	-75	+21	-42	+14	-57
X = P	-28	-76	-23	-43	-22	-58
X = As	-39	-92	-31	-59	-30	-74
X = Sb	-44	-82	-36	-50	-32	-62

BeH_2 are between 75 and 92 kJ mol^{-1}more acidic than expected if the deprotonation would take place at the XH_2 (X = N, P, As, Sb) group. The increase is slightly weaker, from 57 to 74 and from 42 to 59 kJ mol^{-1}, for the complexes with AlH_3 and BH_3 respectively. It is also weaker, but still significant, for the complexes involving the vinyl derivatives, with the only exception, already mentioned above, of vinylamine.

However, to form the insertion-anion ($[R–YH_2–XH_2]^-$), it is first necesarry to form the *insertion-complex* and then, deprotonate this complex. In Fig. 18 is shown the mechanism using as a suitble example the alane adduct of vinylphosphine. Formation of the *insertion-complex* starts from the π-complex **B**, which in general is less stable than the **A** adduct for vinyl derivatives and sligthly more stable for some ethynyl derivatives. From the π-complex **B**, a proton is transferred to the PH_2 group. At the same time two new bonds are formed between Al and each of the C of the vinyl unit, in such a way that a cyclopropane-like ring is formed by Al and the two carbon atoms. This process supposes to cross a very high energy barrier, **TS_BC**, of 225 kJ mol^{-1}. Once cyclic strucure **C** has been reached, the PH_3 moiety starts to interact with the AlH_2, shortening the Al–P distance and lengthening the C–P one so that the AlH_2 subunit inserts in the C–P bond and the *insertion complex* **D** is formed. Again, to pass from **C** to **D** is necessary to cross a high barrier, **TS_CD**. Therefore, formation of the Be, B or As anion implies a complicate mechanism which involve high activation barriers. Hence, in spite of its increased stability, the Be, B, or As anions will most likely not form.

Nevertheless, it is also possible that under the experimental conditions

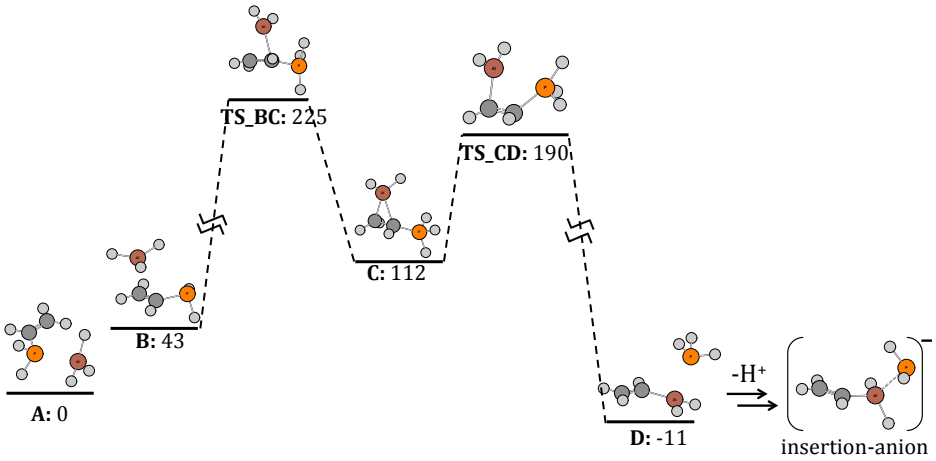


Figure 18: Mechanism to form the insertion-complex **D** starting from the π-complex **B**. Note the high barrier necessary to overcome in this process (225 kJ mol^{-1}). Deprotonation of the *insertion-complex* leads to formation of the most stable anion, [R–YH_2–XH_2]$^-$, insertion-anion.

this rearrangement would be promoted (dynamical factors) and the *insertion complex* formed. The formation of this anion has two major implications: i) the acidity enhancement of the bases considered would be even larger. Taking into account the extra stabilization reported in 7,the complexes between ethynylarsine and ethynylstibine with BeH_2 are predicted to be acids as acidic ($\Delta_{acid}G^0 = 1199$ and 1211 kJ mol^{-1}, respectively) as perchloric acid (($\Delta_{acid}G^0 = 1200 \pm 50$ kJ mol^{-1})! [ref47] Similarly, their complexes with borane, are predicted to be (($\Delta_{acid}G^0 = 1250$ and 1240 kJ mol^{-1}, respectively) more acidic than chloric acid (($\Delta_{acid}G^0 = 1284. \pm 16$ kJ mol^{-1}) [ref48] and sulfuric acid (($\Delta_{acid}G^0 = 1265. \pm 23.$ kJ mol^{-1}). [ref 49]
ii) The adducts would behave as Be, B, or Al Brønsted acids instead of N, P, As or Sb acids. At this point it would be nice to have the experimental gas-phase acidities, to compare the values and thus discrimine between the two isomers for the deprotonated complexes.

Conclusion The intrinsic acidity of the unsaturated vinyl- and ethynylamines, phosphines, arsines and stibines is systematically larger than that of the saturated ethyl analogues, reflecting the larger electronegativity of the vinyl and ethynyl groups with respect to the ethyl one. For the free systems a steadily increase of this intrinsic acidity down the group is observed no matter the nature of the organic moiety to which the XH_2 (X = N, P, As, Sb) acidic site is bound.

Association of the Lewis base with beryllium dihydride, borane or alane leads to a dramatic acidity enhancement due to a much larger stabilization of the deprotonated anion than the neutral molecule. This acidity enhancement, that follows the order: $AlH_3 > BeH_2 > BH_3$, is accompanied, for the unsaturated compounds, by a change in the acidity trends which do not steadily increase down the group, as for the free systems, but present a minimum for both the vinyl- and the ethynyl-phosphine. This unexpected result is due to a much larger increase of the acidity of N containing systems, due to the ability of the N to conjugate with the π-system in the anionic deprotonated species.

For the three Lewis acids and for all the unsaturated amines, phosphines, arsines, and stibines, with the only exception of vinylamine, the most stable anion is obtained by deprotonation at the Lewis acid site (Be, B or Al) rather than in the heteroatom (N, P, As or Sb). It implies that the acidity enhancement reported would be even larger, and most importantly, that the adducts would behave as Be, B or Al acids. However, this deprotonation supposes significant structural rearrangements and it is not likely to occur.

5 Conclusions

Through the use of accurate G4 and DFT calculations we have shown that the association of different Lewis acids with group 15 Lewis bases leads to a dramatic increase of their intrinsic acidities. This acidity enhancement is essentially due to a larger stabilization of the deprotonated species when the Lewis base–Lewis acid bond is formed, because the enhanced donor ability of the anion. Consequently, this is a general phenomenon that shoul be observed for any Lewis base, independently of the nature of its basic site, of its potential acidic site and of the Lewis acid considered. Thus, Lewis adduct formation is a good approach to design stronger Brønsted acids in the gas-phase and even to modulate its gas-phase acidity by appropriately selecting the Lewis acid.

6 Lewis acid deformation

The adducts of group 13 and group 15 elements effectively comprise two distinct units, a group 13 center and a group 15 center, connected by a dative bond, In this Lewis acid/Lewis base adduct, the group 15 species can be considered to provide both electrons for the bond from a lone pair, acting as a two electron donor. Conversely, the group 13 center is electron deficient, with only 6 valence electrons. To complete the valence shell, it accepts two further electrons from the gruop 15 center into a vacant p orbital. This charge donation is particularly important when it concerns the stability of the anionic complexes since the dispersion of the extra electron density into the Lewis acid moiety contributes to the enhanced stablity of the complex with respect to the isolated anionic Lewis bases. An example is shown in Fig. 19. In this figure it can be seen that complex formation supposes a signigicant reduction of the negative charge bear by the vinyl–XH_2 moiety (about 50% for X = N and 75% when X = P).

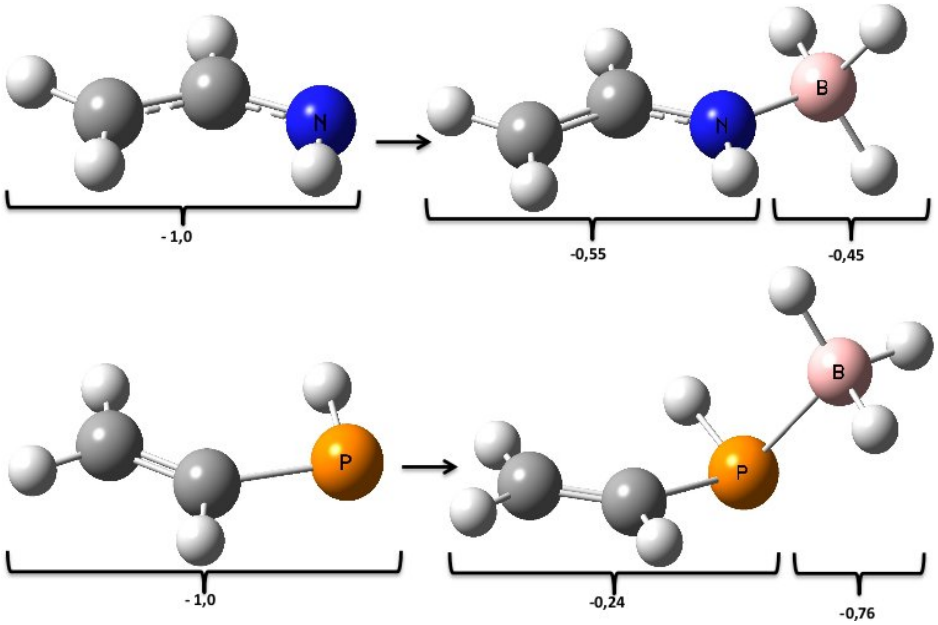


Figure 19: Two examples of the electron distribution within the vinylamine and vinylphosphine molecules upon BH_3 complexation.

Another important common feature observed for all the systems considered is the deformation of the interacting subunits upon complexation. This deformation is usually mirrored in significant changes in their chemical properties. We have already seen the results of adduct formation on the chemical

properties of the Lewis bases. Let us focus now on the structural changes undergone by the Lewis acid moiety when it forms complexes and the effects on its intrinsic properties. We will use the adducts of BH_3 and the last set of bases considered, namely, ethyl-, vinyl- and ethynyl-amine, phosphine, arsine and stibine as a suitable example. Formation of the adduct bond leads to a dramatic distortion of the BH_3 which departs significantly from planarity to produce an approximately tetrahedral geometry. Concomitantly, there is a change in hybridization at boron from approximately sp^2 to sp^3. In order to measure the distortion undergone by BH_3 we will use the dihedral angle defined by the four atoms of the molecule. In figure 20 are plotted the values for this angle in the ethyl-, vinyl- and ethynyl- neutral and anionic complexes.

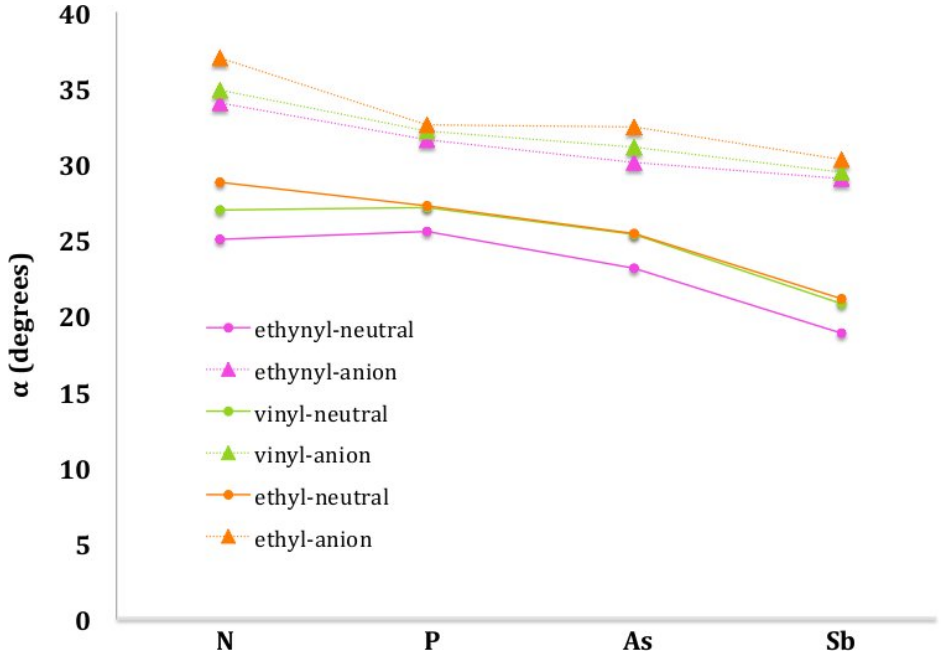


Figure 20: Variation of the BH_3 dihedral angle upon complex formation.

The first conspicuous fact is the notable distortion of BH_3 upon adduct formation, being always larger for the anions than for the neutral complexes (in the free BH_3 equilibrium geometry this angle is zero). The extend of pyramidalization depends however on R substituent or the X heteroatom considered, varying from 19 to 37 degrees. The largest values are always observed for X = N in the anion, while this is only true for the neutral saturated derivatives. The distortion becomes less significant when moving down the group 15.

The important question now is how does the distortion affects to the electronic configuration of BH_3. These electronic changes are crealy mirrored in the Lowest Occupied Molecular Orbital (LUMO) energy of BH_3, which has been plotted in Fig. 21 for the different complexes under survey. It is important to emphasize that in order to obtain valuable information this energies have been computed for the geometry that BH_3 has *within* the complex, rather than for the relaxed BH_3 geometry.

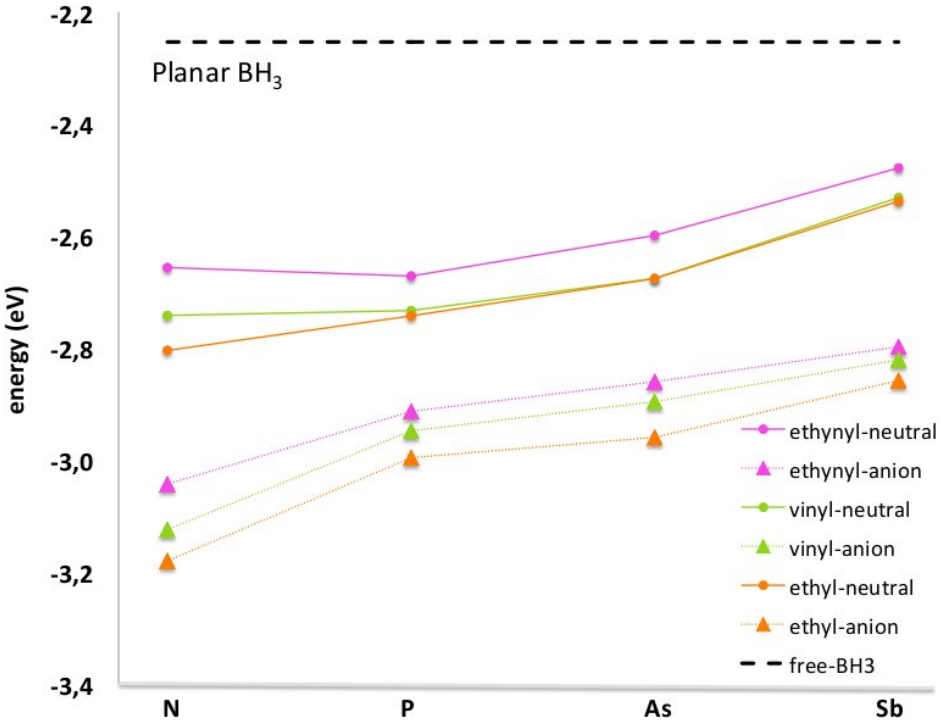


Figure 21: BH_3 LUMO energies within the complex geometry.

Correlation between BH_3 dihedral angle and BH_3 LUMO energy is obvious. Why do the LUMO energies change so dramatically upon complexation and therefore pyramidalization? The bending of BH_3 implies an automatic mixture fo the empty $2p$ orbital and the $2s$ orbital. This mix can not occur in the linear system since these two orbitals belong to different irreducible representaions. Thus, the pure p orbital is stabilized by pyramidalization through the aforementioned mixing with the $2s$ orbital. The larger the deformation, the larger the mix, and the lower the LUMO energy.

It has been proposed that Lewis acidity should be gauged based upon

valence deficiency or the ability of the boron to accept an electron pair.[59,60] Lewis acidity has also been claimed to be correlated to the LUMO (LUMO) level of the Lewis acid.[61] According to the frontier orbital model of chemical reactivity, the lower the LUMO energy of the Lewis acid the stronger would be the interaction with the Lewis base. This is clearly seen in Fig **??** where the LUMO energies of BH_3 computed with the geometry it has within the complex, and the free energies for complexation $\Delta_r G_1^0$ (neutral) and $\Delta_r G_4^0$ (anion) have been plotted. In the case of the unsaturated bases the relationship is not so clear since conjugation of the X lone pair with the π-system of the molecule, significant for the anionic species and for neutral amines implies and extra stabilization present on the values of the free energies but, of course, not mirrored on the LUMO energies of the BH_3.

Therefore, the structural changes undergone by the Lewis acid moiety upon complexation trigger a rearangement of the electron distribution that implies a decrease on the BH_3 LUMO energy. Consequently, the intrinsic acidity of the Lewis acid is strongly modified upon adduct formation.

The role of Lewis acid deformation on the stability trends of the complexes formed between ammonia and $BeH_{2-n}X_n$ (X = F, Cl, Br; n $\leq$ 2) Lewis acids its showed in the paper entitled *The importance of deformation on the strength of beryllium bonds* that is attached in the Appendix.

7 Radicales

redINTRO A LOS RADICALES. LUEGO VER SI PONGO EL PAPER EN UN APNDICE O DIGO ALGO AQUI

Cleavage of the X–H bond can occur heterolytically, as has been described until now, or, alternatively, homolytically to yield two uncharged radical species: H· and RX·.

Amine–Borane radicals. As part of the reactivity of amine–borane and phosphine–boranes there are the radical reactions. Amine–borane adducts have been shown to react rapidly with radical species, which abstract a hydrogen atom from boron to produce transient amine–boryl radicals (see Fig. 23).

The seminal research in this area ws initiated in the mid 1980s by Roberts and co-workers, who performed a large number of electron spin resonance

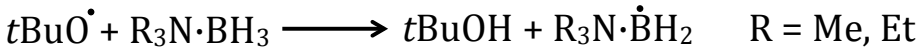


Figure 22: Hydrogen abstraction from trialkylamine–boranes by *t*-butoxylradicals.

(ESR) based studies on radical species resulting from hydrogen abstraction from *N*,*N*,*N*-trialkyl–[62,63] and *N*,*N*-dialkulamine–boranes,[64] and subsequently ammonia–borane.[65] Various reactivities of these amine–boryl radicals were demonstrated, all resulting in the quenching of the boryl radical. The reactivity is fundamentally similar to that of carbon centered radicals, with consecutive reactions occuring via (a) self-reaction, (b) abstraction of halide radicals from alkyl halides or H· radicals from other amine-boranes, and (c) β-scission. The reactivity was strongly dependent on the nature and substitution of the amine, and reaction conditions.

Of particular interest was reaction **??**, wherein a *N*,*N*-dimethylamine–boryl radical reacted with a molecule of its precursor *N*,*N*-dimethylamine–borane to form *N*,*N*-dimethylaminyl–borane, a nitrogen centered radical. This reaction was a simple conversion between the kinetic product, where the radical was located at boron, and the thermodynamic product, where the radical was located at nitrogen.[64] Attack at boron of the alkoxyl radical to produce the *N*,*N*-dimethylamine–boryl radicals was initially favored because of polar effects operating in the transition state of hydrogen abstraction: the electrophilic alkoxyl radical preferred to attack at the electron-rich borane moiety rather than the electron-poor amine component.

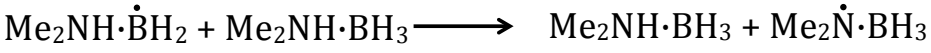


Figure 23: Formation of *N*,*N*-dimethylaminyl–borane, a nitrogen centered radical.

A prominent use developed for amine–boryl radicals is the field of polarity reversal catalysts in radical reactions.[66]

Phosphine–borane complexes in radical chemistry. *In situ* generate phosphine–boryl radicals have been introduced by Barton as deoxygenating reagents for xanthates. prepared from sterically hindered secondary alcohols. Roberts and co-workers had shown earlier that phosphine–boryl radicals re-

act with butyl iodide to abstract an I· radical. The resulting *n*-butyl radical could then add to an olefin and abstract a further hydrogen atom to give the alkane.[67,68]

It was demonstrated that such phosphine–boranes could be employed in a Barton–McCombie reaction in place of the more toxic trialkyltinhydrides as hydrogen donors.[69] Since then, this reaction has found application in organic synthesis, for example in the synthesis of a natural product analogue of tyrsiferol[70,71] and others.[72–74]

8 Conclusiones

A lo largo de esta primera parte se han expuesto los valores calculados tericamente (y en algunos casos tambin experimentalmente) para una serie de bases de Lewis y los complejos formados con una serie de cidos de Lewis. Se observa sistemticamente un notable aumento de la acidez intrnseca de dichas bases, que llega a ser en algunos de 46 rdenes de magnitud en trminos de la constante de acidez. Este significante aumento de acidez se explica debido a la mayor estabilizacin de la forma neutra de la base al formarse en complejo en relacin con la correspondiente estabilizacin de la forma neutra.

PART II

1 Introduction

Doubly charged molecular ions, formed by association of a neutral base with a doubly charged metal ion, are of great importance in chemical and biochemical processes, boh in solution and in the gas phase.[75] Indeed, they are fundamental in gas-phase ion chemistry and in mass spectrometry in spite of their short life-time.[?] They also play important roles in photochemicall processes,[?,?] as well as in astrochemsitry and the chemistry of the atmospheres.[?] It has also been postulated that they may have played some role in the origin of life,[?] since doubly charged metal ions may have induced the precipitation of nucleic acid molecules. Newvertheless, these multiply charged ions were, for a long time oddities in the gas phase due to the difficulties in generating and stabilizing them. The so-called *intrinsic reactivity* is of great importance since the absence of interactions with a solvent can result in very different reactivity patterns; in many cases it allows for a better understanding of chemical bonding.

With the advent of electrospray ionization techniques in 1990,[?] generation of doubly-charged ions in the gas-phase from aqueous solutions became feasible and therefore the interest in gas-phase reactions between di-cations and neutral molecules has grown significantly. Many theoretical studies have been performed to understand the structures and relative stabilities of such ion-molecule complexes, providing detailed information on the nature of their interactions and binding energies.[?] However, information about their unimolecular reactivity is much scarcer and fragmentary. One of the main reasons is that doubly- or multiply-charged species are rare in the gas-phase because many of these species are either thermochemically or kinetically unstable.[75] This is indeed the case when they are the result of the association of a doubly (or multiply) charged transtition metal ion (M^{2+}) with an organic base (B). The system undergoes a spontaneous deprotonation of the base leading to the monocation $[M(B\text{-}H)]^+$, the species experimentally observed.[?]

However, this is not the case with alkaline-earth dications such as Ca^{2+} and Sr^{2+}, since $[CaB]^{2+}$ and $[SrB]^{2+}$ are stable and detectable in the gas phase.[?] This opened the possibility of studying for the first time the unimolecular reactivity of complexes formed by these metal dications with different neutral bases such as urea, glycine, thiourea, selenourea, uracil and its thio derivatives.[?] All these studies show that there is a competition between Coulomb explosion and neutral loss fragmentation.

Different techniques can be used for activating ions, which differ by the energy range employed, the instrument and the activation mechanisms.[?] Notable among them is collision induced dissociation (CID), where the ion is made to collide with an inert gas at low collision energies, in such a way that part of the kinetic energy is transferred to the molecular ion as vibrational and rotational internal energy. CID has been extensively used to study ion structure and is one of the most commonly ion activation methods.[?,?] Nevertheless, there are many details of gas-phase dissociation mechanisms associated with CID that are still unclear. In this realm, theoretical calculations can help to understand the CID processes and in particular provide an atomistic description of the mechanisms and pathways leading to the final fragments.

Recently, Eizaguirre *et al.* studied the interactions and dissociation patterns of $[M(formamide)]^{2+}$ (M = Ca, Sr) complexes.[?,?] The study combined collision induced dissociation (CID) experiments and high-level DFT calculations to analyze the topology of the potential energy surface (PES). A mechanism leading to the different fragments observed in the experiment was proposed based on the computed PESs. The experimental CID spectra for formamide-M^{2+}, M= Ca, Sr; are shown in Fig. 24 and the corresponding PESs proposed are shown in Fig. 25 (M = Ca) and 26 (M = Sr). Although it is possible to extract some conclusions about the reactivity from the topology of the PES, this constitutes only a first rough approach, and there are still open questions. For instance, whereas both molecules display a PES quite similar, the experimental CID spectra for M = Ca shows the presence of fragmentation pathways nonexistent when M = Sr. Some other features of the CID spectra, as the presence of a very intense peak corresponding to the bare metal dication, M^{2+}, cannot be accounted for based solely on the topology of the PES, since this fragmentation corresponds to the most endothermic exit channel. Thus, we decided to further study CID reactivity of formamide-Ca^{2+} and formamide-Sr^{2+} dications by means of theoretical approaches different from the one previously used by Eizaguirre *et al.*

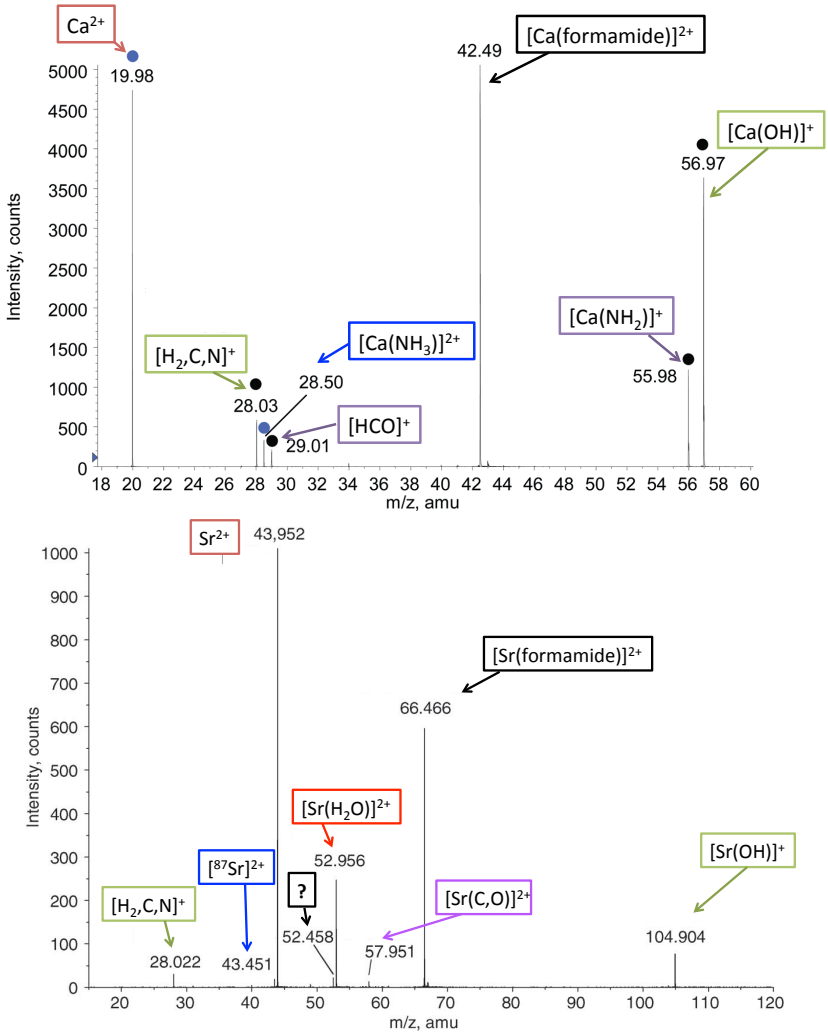


Figure 24: Top: low-energy CID spectrum of the formamide-Ca^{2+} complex recorded with a collision energy of 11eV (laboratory frame). Bottom: low-energy CID spectrum of formamide-$^{87}Sr^{2+}$ recorded at a collision energy of 14eV (laboratory frame).

From the theoretical point of view, statistical models such as transition state theory (TST) and Rice-Ramsperger-Kassel-Marcus (RRKM) theory[?] have been (and are) extensively used to describe the kinetics of the unimolecular reaction undergone by the ion after the collisional excitation.[?,?,?,?] However, evidence of non-statistical dynamics, also called non-RRKM, in CID has been reported both in experiments[?] and simulations,[?] as well as the importance of shattering mechanisms to understand CID in protonated

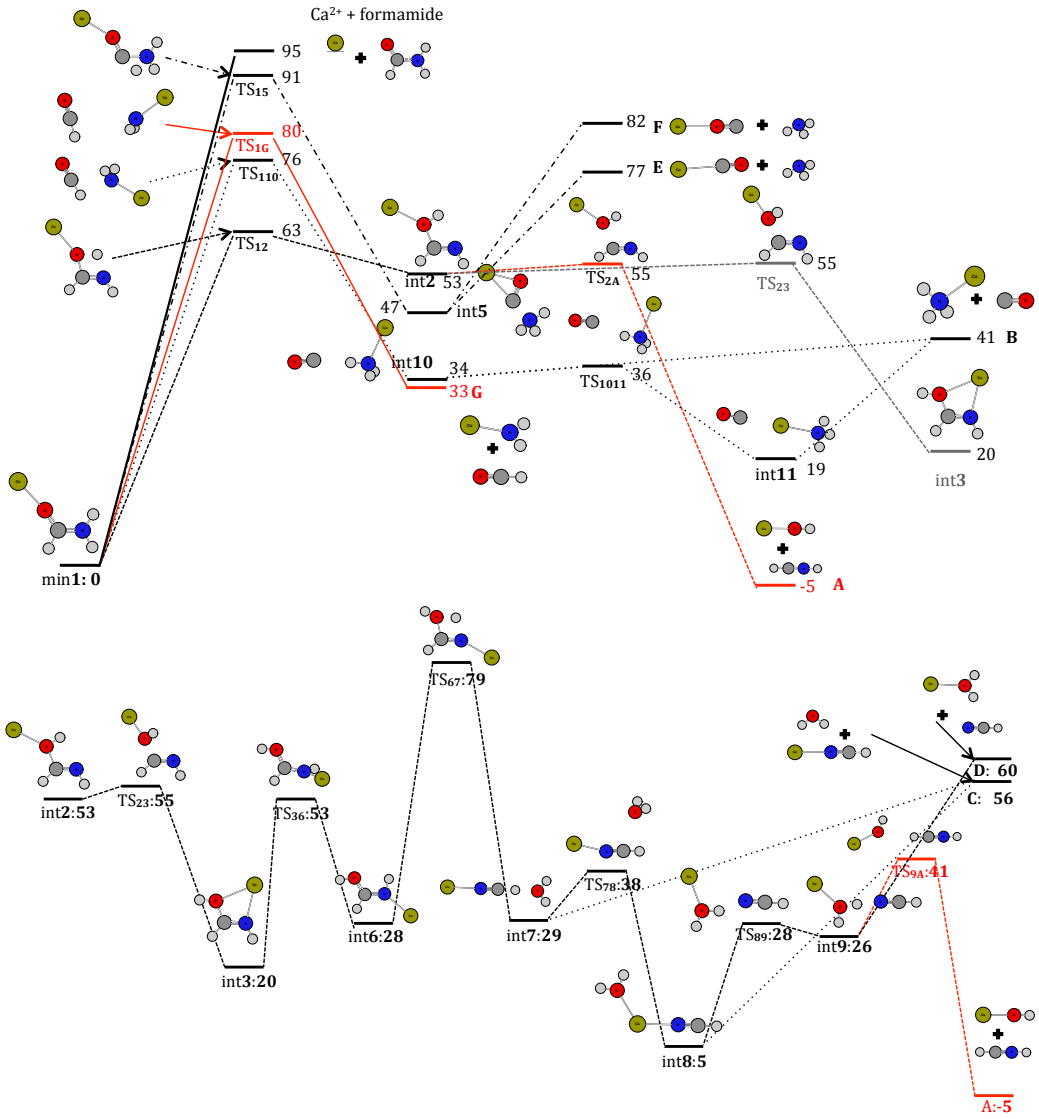


Figure 25: Top: formamide-Ca^{2+} energy profile of the different reaction mechanims with origin in the global minimum **1** and (bottom) local minimum **2**. All values in kcal mol^{-1}.

amino acids and peptides.[?]

Two limiting fragmentation mechanisms can be described to explain CID reactivity. One in which the collision locally activates one (or few) vibrational mode(s), and the fragmentation occurs within one vibrational period before intramolecular vibrational energy redistribution (IVR) could take place. In

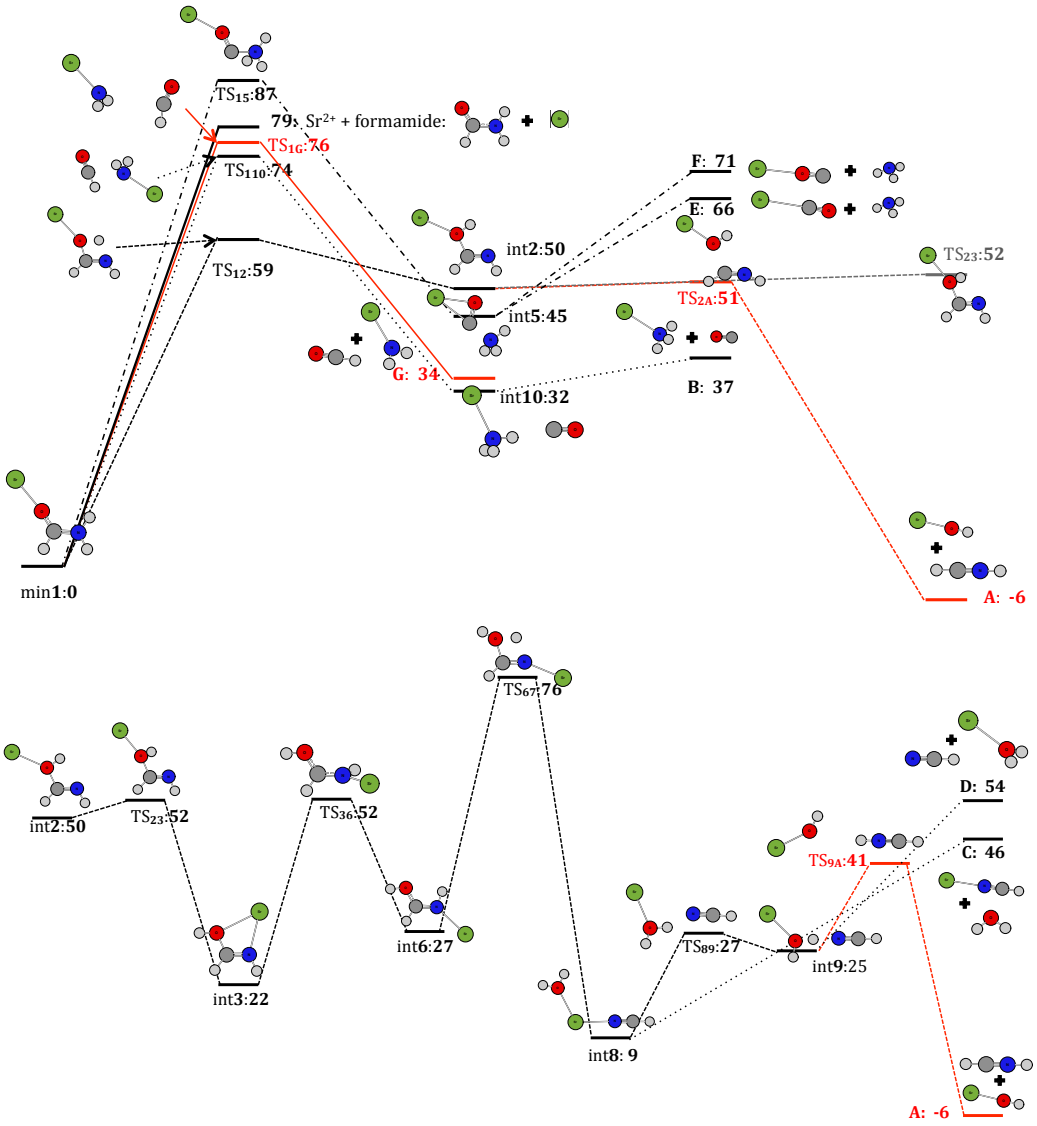


Figure 26: Top: formamide-$^{87}Sr^{2+}$ energy profile of the different reaction mechanisms with origin in the global minimum **1** and (bottom) local minimum **2** (bottom). All vales in kcal mol^{-1}.

the second mechanism, the translational energy transferred to the molecule is redistributed among its vibrational and rotational modes. If the energy in a vibrational mode is higher than the energy required for breaking this bond the molecule eventually dissociates. Reactions taking place via the latter mechanism can be accounted for through statistical theories such as RRKM, while in the case of the first mechanism, a pure dynamical picture where the

reaction is faster than IVR, the reactivity can only be understood by means of dynamics simulations. Therefore, an important question concerning CID unimolecular reactivity is whether the mechanisms leading to the different fragmentation pathways are statistical or not.

Thus, the purpose of this part of the thesis is twofold: on the one hand to investigate and characterize the fragmentation mechanisms on CID experiments and, at the same time, provide an explanation to the differences observed in the CID spectra of $[Ca(formamide)]^{2+}$ and $[Sr(formamide)]^{2+}$ dications. With this aim we performed chemical dynamics simulations and compared the results with the predictions obtained using RRKM statistical theory (short-time regime). Furthermore, by coupling the vibrational and rotational energy distributions obtained from chemical dynamics simulations of non-reactive trajectories with RRKM rate constants, $k(E)$, it is possible to model reactivity in the long-time scale, an important issue that is difficult to directly obtain only from chemical dynamics simulations. Thus, the combination of both methods provides a multi-scale approach allowing to bridge the gap between the different time-scales involved in CID experiments: short-time non-statistical reactivity together with the long-time statistical reactivity.

2 Method assessment

Such calculations require a reliable description of the potential energy surface as well as reliable predictions of molecular properties, and this can be attained through the use of *ab initio* methods. CID experiments are now routinely applied to small and large molecules such as peptides and proteins,[?] and combining CID with RRKM in these latter cases requires electronic representations computationally less expensive than *ab initio* methods. A similar situation occurs when coupling RRKM and direct dynamics simulations of the fragmentation.[?] This calls for resorting to less computationally demanding theoretical approaches, such as density functional theory (DFT).[?] To be sure about the reliability of the results obtained *via* DFT it is, however, necessary to carry out an assessment in order to choose the most appropriate functional.

Thus, one of the aims of this section is to test a variety of DFT methods, as well as the MP2 wave function method against benchmark CCSD(T) results for the two aforementioned systems, using a small basis set (6-31G(d)).

This would allow us to validate the accuracy of non-expensive methods that might be used for computationally expensive treatments such as dynamics simulations or bigger systems like oligopeptides. The geometries and energies will be obtained with 21 different DFT approaches. These methods will also be assessed from the kinetic point of view, a rather unusual approach providing an original way to assess DFT functionals also on kinetics rather than simply on energetics and/or frequencies.

Preliminary assessment A wide window of functionals combined with a small bases set were evaluated in a preliminary assessment ("low cost" methods). However, caution must be taken when selecting a DFT method to use for a specific problem or a specific system because often a DFT method that correctly predicts certain properties, such as geometries or binding energies, will prove to be much less accurate for the computation of other properties, such as barrier heights of conformational energy differences. On the other hand, those energies are the energetic properties governing the different reactive pathways. Therefore, the crucial point is whether this "low cost" approach performs well enough when treating di-cationic systems. Hence, in this assessment, we considered energetics corresponding to the principal reaction mechanisms: M^{2+}–O dissociation energy to obtain formamide + M^{2+}, as a representative example of neutral loss, and the barrier height for the fragmentation of the global minimum (min**1**) into $[M(NH_2)]^+ + [HCO]^+$, as an example of Coulomb explosion. The goal is to find the method that performs better for both processes (metal–ligand bond energy and barrier height). For this assessment we will use CCSD(T) single-point calculations on DFT optimized geometries as reference. We evaluated 21 functionals using the 6-31G(d) basis set plus three functionlas (BLYP, G96LYP) and B3LYP) using the 6-31+G(d,p) basis set. In all cases the basis set for Sr is LANL2DZ. For more detailed information about the functionals and basis set employed see section the methodology section. The results are shown in Tables 8 and 9 for M = Ca and M = Sr, respectively.

Table 8: Ca^{2+}–O dissociation energy (NL) and Coulomb explosion ($[Ca(NH_2)]^+$ + $[HCO]^+$) energy barrier (CE) computed with different DFT functionals, all with the 6-31G(d) basis set but when explicitly written. The third column shows the energy difference between Ca^{2+}–O dissociation energy and Coulomb explosion energy barrier (gap). The fourth column is the same energy difference without taking into account the ZPVE correction (gap*). The first row corresponds to the benchmark method. All values are in kcal mol^{-1}.

formamide-Ca^{2+}				
	NL	CE	Gap	Gap*
CCSD(T)/cc-pWCVTZ	**90.76**	**77,45**	**13.31**	**10.10**
BLYP	88.02	81.51	6.51	2.95
BLYP/6-31+G(d,p)	82.74	78.77	3.97	0.38
G96LYP	84.30	82.22	2.09	-1.57
VSXC	86.42	85.34	1.08	-2.60
G96LYP/6-31+G(d,p)	80.19	79.71	0.48	-3.10
PBEPBE[a)]	87.50	88.10	-0.60	-4.23
MP2	86.24	87.62	-1.38	-4.46
O3LYP	85.21	88.66	-3.45	-6.70
X3LYP[a)]	89.74	93.42	-3.69	-6.94
B3LYP	88.86	92.58	-3.72	-6.94
B3LYP/6-31+G(d,p)	84.58	88.80	-4.22	-7.51
TPSS	87.14	92.32	-5.18	-8.55
M06-2X[a)]	89.22	95.53	-6.30	-9.05
M05	87.22	93.90	-6.68	-9.67
M05-2X	90.04	99.18	-9.13	-11.86
M06[a)]	86.21	95.39	-9.17	-12.27
BH&HLYP	90.18	99.68	-9.50	-12.57
B972	85.60	95.23	-9.63	-12.85
B3PW91	86.23	96.05	-9.81	-13.11
M06L[a)]	82.73	92.99	-10.26	-13.91
MPW1PW91	87.42	98.17	-10.75	-13.95
PBE1PBE[a)]	87.87	98.74	-10.86	-14.17
B3P86	87.27	98.24	-10.97	-14.30
BMK	86.59	99.49	-12.90	-15.95
BH&H	95.56	110.00	-14.44	-17.32

[a] ZPVE wsa not corrected by the scale factor because it was not available.

Table 9: Sr^{2+}–O dissociation energy (NL) and Coulomb explosion ($[Sr(NH_2)]^+$ + $[HCO]^+$) energy barrier (CE) computed with different DFT functionals, all with the 6-31G(d) basis set but when explicitly written. The third column shows the energy difference between Sr^{2+}–O dissociation energy and Coulomb explosion energy barrier (gap). The fourth column is the same energy difference without taking into account the ZPVE correction (gap*). The first row corresponds to the benchmark method. All values are in kcal mol^{-1}.

formamide-Sr^{2+}				
	NL	CE	Gap	Gap*
CCSD(T)/6-311+G(3df,2p)	**78.97**	**86.95**	**-7.98**	**-11.60**
G96LYP/6-31+G(d,p)	73.28	81.57	-8.30	-11.91
BLYP/6-31+G(d,p)	63.82	85.49	-21.67	-25.79
BLYP	63.92	87.52	-23.61	-27.86
VSXC	65.98	91.38	-25.39	-29.18
G96LYP	60.68	87.97	-27.29	-31.31
PBEPBE[a)]	64.55	94.35	-29.80	-33.85
B3LYP/6-31+G(d,p)	66.58	98.09	-31.50	-35.26
O3LYP	64.07	97.17	-33.11	-37.29
TPSS	66.32	100.40	-34.07	-38.24
M06L[a)]	65.12	99.54	-34.42	-38.66
B3LYP	66.17	101.90	-35.72	-39.68
X3LYP[a)]	66.93	103.13	-36.20	-40.15
06-2X[a)]	69.08	105.52	-36.44	-38.99
M05	66.02	104.81	-38.80	-42.11
B972	65.56	104.85	-39.29	-43.20
M06[a)]	65.59	105.38	-39.79	-43.14
MP2	63.01	102.86	-39.85	-43.34
B3PW91	65.19	105.31	-40.12	-44.11
BH&HLYP	69.58	110.35	-40.77	-43.69
B3P86	66.21	107.48	-41.27	-45.26
MPW1PW91	66.75	108,35	-41.60	-45.38
PBE1PBE[a)]	66.98	109.07	-42.09	-45.96
M05-2X	69.15	112.44	-43.29	-46.48
BMK	65.91	111.91	-46.00	-48.68
BH&H	73.02	119.79	-46.77	-49.75

[a] ZPVE wsa not corrected by the scale factor because it was not available.

From the values in Table 8 we observe that for M = Ca only five methods give the correct qualitative result showing that as predicted by the reference CCSD(T) calculations the neutral loss is higher in energy than teh activation barrier for the Coulomb explosion (positive values). In the case of Sr containing systems (Table 9) all the functionals reproduce CCSD(T) results qualitatively (same sign). The best performing method, as might be expected, depends on the metal: when M = Ca the best performing model is BLYP/6-31G(d) while when M = Sr the best one is G96LYP/6-31+G(d,p); it gives almost the same values as the reference. This latter result is in agreement with previous assessments,[?] and this level is the one used in geometry optimization for the Sr containing system. Therefore, we have selceted these methods for a subsequent assessment evaluating the whole PES, comparing geometries and relative energies *vide infra*. It should be noted that in general all the functionals investigated perform well for the description of the NL dissociation energy, so the differences in correctly reproducing the gap aminly arise from the significant differences in the estimation of the Coulomb explosion barrier. Indeed, the functionals like BLYP that do not include Hartree-Fock (HF) exchange perform better than the hybrid ones. Actually, it can be observed that the agreeement with CCSD(T)-reference value decreases as teh amount of Hartree-Fock exchange increases, being worse for BH&HLYP than for B3LYP. In general, the inclusion of HF exchange introduces a partial self-interaction correction, but at the same time it removes the non-dynamic correlation effects described by the GGA functionals.[?,?] This actually may explain the good performance of methods like BLYP for the particular case of the Coulomb explosion barrier, where static correlation must be important taking into account that, at the barrier, the wavefunction must be the mixture of covalent and ionic components. However, the good performance could be due to cancellation of errors affecting both estimated energies.

Since one of our objectives is to compare the formamide-Ca^{2+} and formamide-Sr^{2+} reactions, it is advisable to use the same theoretical model to reproduce both PESs. Hence, for this purpose we have chosen the G96LYP/6-31G(d) approach, which is among the best performing for both systems. For the sake of completeness, we will also test the heavily used B3LYP functional and the MP2 *ab initio* method using in all cases a 6-31G(d) basis set. In this assessment we have used as reference the potential energy surface obtained by Eizaguirre *et al.* for formamide-Ca^{2+} and formamide-Sr^{2+}.[?,?]

Geometry assessment. A comparison of the geometries obtained with the aforementioned four methods with those reported previously[?,?] shows rather small differences for both metals (Ca and Sr). This is very well reflected by the small errors affecting the rotational constants A, B. and C (see Fig. 27), which are sensitive criteria when comparing different geometries. All the values are given in Tables A2.19–A2.24 in the Appendix 2.

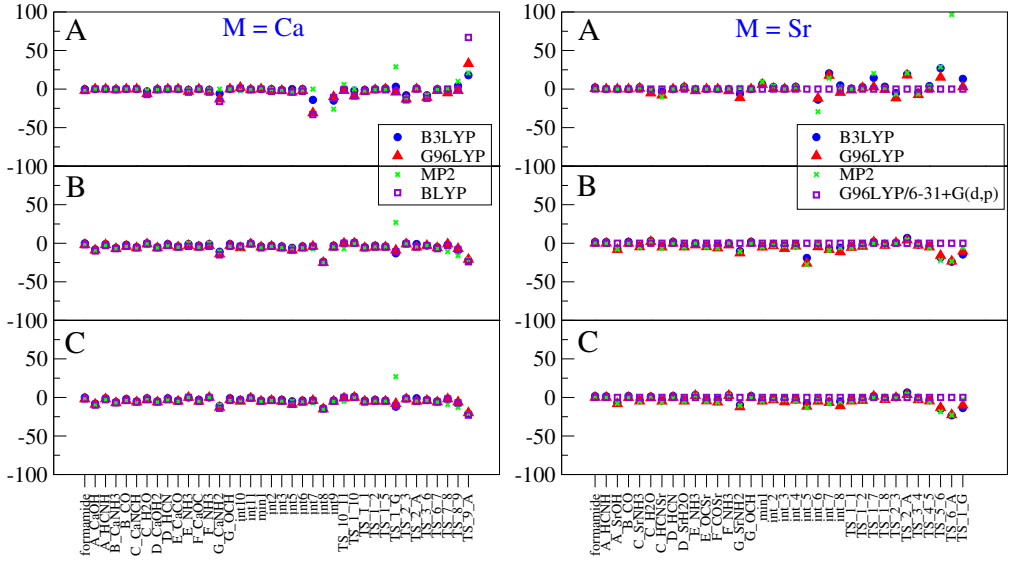


Figure 27: Relative error (%) derivations for rotational constants of (left) formamide-Ca^{2+} computed at B3LYP/6-31G(d) (blue circles), G96LYP/6-31G(d) (red triangles), MP2/6-31G(d) (green crosses), and BLYP/6-31G(d) (purple squares) and (right) formamide-Sr^{2+} computed at B3LYP/6-31G(d) (blue circles), G96LYP/6-31G(d) (red triangles), MP2/6-31G(d) (green crosses), and G96LYP/6-31+G(d,p) (purple squares).

The results show that globally there are no significant differences in the geometries optimized with the low-level approaches, although for some structures the magnitude of the relative error is slightly large. We will briefly comment on the structures with teh most significant deviations, namely int**8** and TS_1_G. Tables with all relative error values are given in Appendix 2 (Tables A2.19–A2.24).

The structure with highest errors is int**8** for both metals. These errors in three rotational constants arise from differences in the oxygen-metal-nitrogen angle, which is about 180° with the trial methods, and 120° in the reference

structure (see Fig. 28a). As expected, the rotational constant A is the most sensitive to structural changes. The TS_1_G strucure is well described with the three DFT functionals while MP2 overestimates the three rotational constants by about 30% for M = Ca and the rotational constant A (268% relative error) for M = Sr. These discrepancies are mainly due to differences in the relative orientation of the two deparing fragments (Fig. 28b). Another structure with remarkably relative errors is TS_9_A, in particular when M = Sr, again due to different relative orientations of the departing fragments.

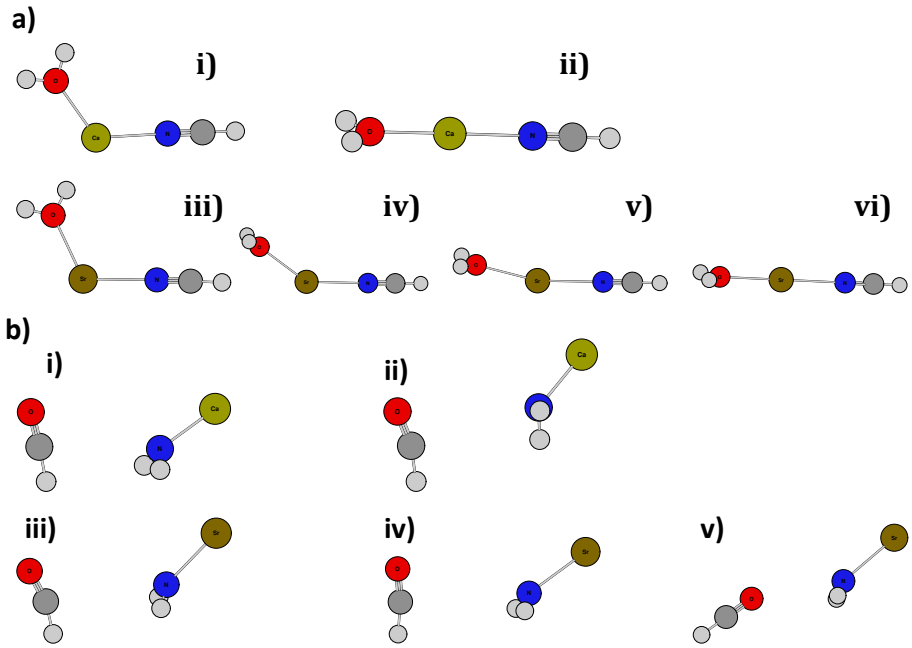


Figure 28: (a) Int**8** structure for M = Ca optimized with (i) B3LYP/cc-pWCVTZ, (ii) BLYP/6-31G(d) and for M = Sr optimized with (iii) G96LYP/6-31+G(d,p), (iv) B3LYP/6-31G(d), (v) G96LYP/6-31G(d) and (vi) MP2/6-31G(d). (b) TS_1_G structure for M = Ca optimized with (i) B3LYP/-ccpWCVTZ, (ii) BLYP/6-31G(d) and for M = Sr optimized with (iii) G96LYP/6-31+G(d,p), (iv) B3LYP/6-31G(d) and (v) MP2/6-31G(d).

In summary, the only structures that show more significant errors with respect to the reference arise from the different orientation of subunits that are weakly bound. The PES, computed at the reference level. together with the structure of the corresponding stationary points are shown in Fig. 25 and 26.

Energy assessment. Let us focus now on the performance of the low level approaches when dealing with energies. Relative energies were evaluated for each method by subtracting from the energy of the corresponding structure (the sum of the energy of the fragments for the exit channels), the energy of the global minimum, min**1**; including the corresponding ZPVE corrections. The absolute error and relative error values for all the structures are given in Tables A2.25 and A2.26 in Appendix 2. Fig. 29 shows the relative error for M = Ca (a) and M = Sr (b). In general the agreement is quite good and the trends are the same for the four methods.

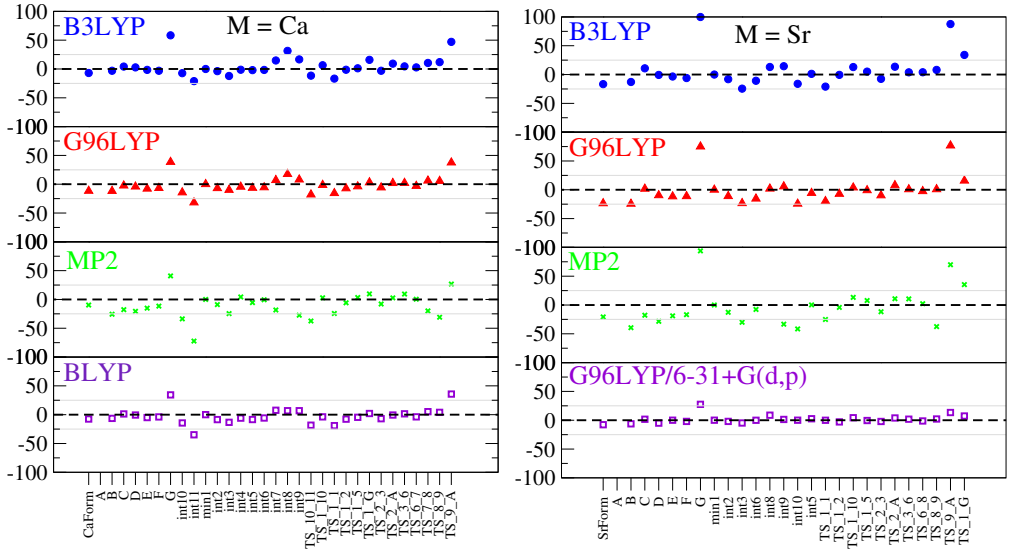


Figure 29: Relative errors in (a) formamide-Ca^{2+} relative energies computed with B3LYP/6-31G(d) (blue circles), G96LYP/6-31G(d) (red triangles), MP2/6-31G(d) (green crosses), and BLYP/6-31G(d) (purple squares) and (b) formamide-Sr^{2+} relative energies with B3LYP/6-31G(d) (blue circles), G96LYP/6-31G(d) (red triangles), MP2/6-31G(d) (green crosses), and G96LYP/6-31+G(d,p) (purple squares).

The root-mean-square error or deviation (RMSD) of the values obtained with the low-level approach from the values computed at high-level, is reported in Table 10. The normalized RMSD (NRMSD = RMSD/$(b_{max} - b_{min})$ where b stands for reference values) is also reported because it is useful to compare how the performance of each method varies when the cationic metal changes from Ca^{2+} to Sr^{2+}.

It can be seen that for M = Ca all methods perform almost equally wel, whereas for M = Sr, G96LYP/6-31+G(d,p) seems to be the best option since

Table 10: Root mean-square-deviation (RMSD) in kcal mol^{-1} and normalized RMSD (NRMSD).

	Formamide-Ca^{2+}				Formamide-Sr^{2+}			
	B3LYP[a]	G96LYP[a]	MP2[a]	BLYP[a]	B3LYP[a]	G96LYP[a]	MP2[a]	G96LYP[b]
RMSD	7.5	6.4	8.6	6.0	14.5	12.5	14.7	3.2
RMSD	8%	6%	9%	6%	16%	13%	16%	3%

[a] 6-31G(d). [b] 6-31+G(d,p)

it shows the lowest RMSD. Let us reacll that G96LYP/6-31+G(d,p) is the method used in the geometry optimization of the benchmark values for M = Sr. These results are in agreement with those obtained in the first assessment discussed above. The next best performance is obtained using G96LYP/6-31G(d) for both systems. In general, there is a better agreement between low-level and high-level when the cation is Ca^{2+}.

It is worth noting that whereas the first assessment of formamide-Ca^{2+} (see Table 8) shows a better performance for MP2/6-31G(d) than for B3LYP/6-31G(d), this is not the case when comparing the whole PES, where B3LYP has a slightly better performance than MP2.

Kinetic assessment. A further step in our assessment needs to include the kinetic behaviour. To this end, we computed the RRKM rate constants, $K(E)$, corresponding to the different reactions shown in Fig. **??**. This figure shows the kinetic schemes associated with the PES for formamide-M^{2+} (M = Ca, Sr) unimolecular reactions (see Fig. **??** and **??**). For the exit channels there is no backward reaction because in CID experiments there is no equilibrium between fragments and parent ions. To compute the values for $k(E)$ we emplyed the same four methods used before for the geometry and energy assessments. The evaluation of the rate constant will imply an indirect assessment of three properties: energy barriers, geometries, and harmonic frequencies, needed to compute the RRKM rate constants.

From now on we will use half-life times, $t_{1/2}$, instead of rate constants, k. This quantity is proportional to the reaction rate constant $t_{1/2} = -(\ln 0.5)/k(E)$ and gives us an idea of how fast a reaction occurs. In order to compare the performance of the different methods, the curves for $t_{1/2}$ *vs.* E where shifted

in the x-axis at a quantitiy equal to the activation barrier energy, E_{act}. Thus E stands for the internal energy available to react, *i.e.*, internal energy over the activation barrier. A graphical example is shown in Fig. 30. The average of the relative errors is plotted in Fig. 31.

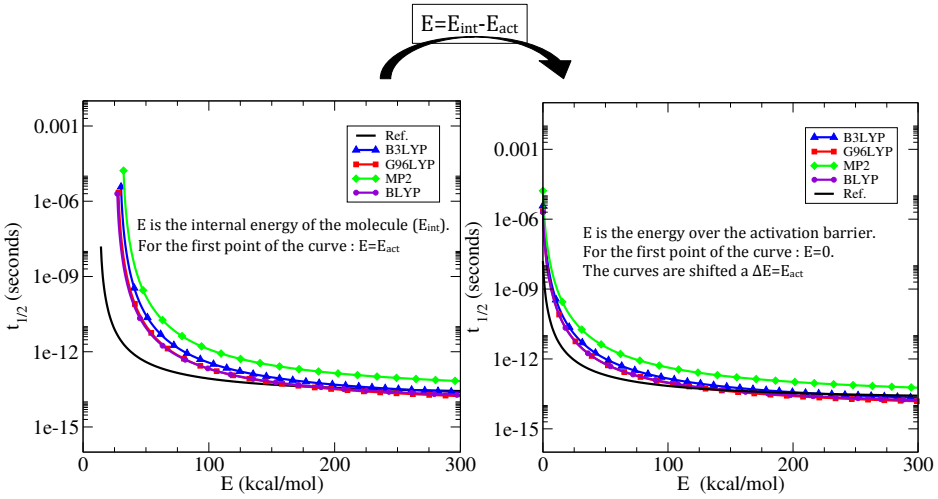


Figure 30: Graphical example showing the x-axis shift performed in the $t_{1/2}(E)$ curves in order to compare the different methods tested.

It should be noted that $t_{1/2}$ is a very sensitive quantity, especially at small E, just over the energy barrier, so not surprisingly the relative errors are large in this region. As the energy increases the errors decrease and the curves tend to converge to a given value. In general, trends are similar for both metals. The MP2 method is clearly th emost unfavourable with very high relative errors. B3LYP and G96LYP perform much better than MP2, but still the relative errors are significant. At small energies G96LYP is better than B3LYP but as the energy increases the order changes, with B3LYP becoming slightly more preferable. The best performance is observed for BLYP when M = Ca, and G96LYP/6-31+G(d,p) for M = Sr.

Conclusions Among all the approaches considered, it was found that G96LYP/6-31G(d) is the best compromise to reproduce PES and kinetics obtained from higher level calculations, for both formamide-Ca^{2+} (B3LYP/cc-pWCVTZ) and formamide-Sr^{2+} (G96LYP/6-311+G(d,p)). Thus, this level of theory will be used to compare the reactivity of both systems.

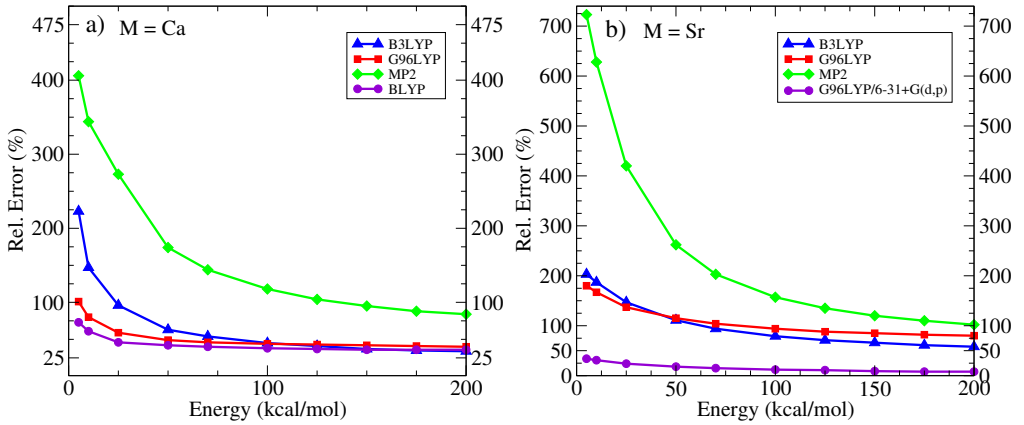


Figure 31: Relative errors in (a) formamide-Ca^{2+} relative energies computed with B3LYP/6-31G(d) (blue circles), G96LYP/6-31G(d) (red triangles), MP2/6-31G(d) (green crosses), and BLYP/6-31G(d) (purple squares) and (b) formamide-Sr^{2+} relative energies with B3LYP/6-31G(d) (blue circles), G96LYP/6-31G(d) (red triangles), MP2/6-31G(d) (green crosses), and G96LYP/6-31+G(d,p) (purple squares).

3 CID of formamide-M^{2+}

The purpose of this section is twofold: on the one hand to investigate and characterize the fragmentation mechanisms on CID experiments and, at the same time, provide an explanation to the differences observed in the CID spectra of formamide-Ca^{2+} and formamide-Sr^{2+} dications. With this aim we compare the results of chemical dynamics simulations with with the predictions obtained using RRKM statistical theory (short-time regime). Furthermore, by coupling the vibrational and rotational energy distributions obtained from chemical dynamics simulations of non-reactive trajectories with RRKM rate constants, $k(E)$, it is possible to model reactivity on the long-time scale, an important issue that is difficult to directly obtain only from chemical dynamics simulations. Thus, the combination of both methods provides a multi-scale approach allowing to bridge the gap between the different time-scales involved in CID experiments: short-time non-statistical reactivity together with the long-time statistical reactivity.

3.1 Direct dynamics simulations

Direct dynamics simulations set-up. As explained in the methods section, the molecules nuclei positions, q_i and momenta p_i, evolve on the Born-

Oppenheimer potential energy surface obtained by solving the time independent Schrödinger equation at each configuration. For this purpose and based on the previous assessment work, we used the G96LYP[?] and BLYP[?,?] functionals, both with the 6-31G(d) basis set for M = Ca and the G96LYP functional with 6-31G(d) basis set for one set of trajectories and 6-31+G(d,p) for other set for M = Sr.

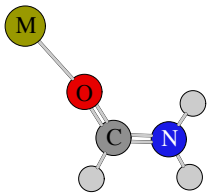


Figure 32

In our approach we just modelled a single collision, which is a good approximation when the gas pressure is very low. For the initial conditions we used the minimum energy structure (min1) of formamide-M^{2+} (M = Ca, Sr) (see Fig. 32).

Since the electrospray ionization source used in the CID experiments is not thermalized we choose an initial temperature of 300 K for the ions as it is usually done in similar studies.[76–79] Energies for the normal modes of vibration were selected from a 300 K Boltzmann distribution. The resulting normal modes energies were partitioned between kinetic and potential energy by choosing a random phase for each normal mode. Rotational energy and angular momentum for the polyatomic molecule were selected by assuming separability of vibrational and rotational motion. Thus initial rotational conditions are obtained by assuming a thermal partitioning of RT/2 about each internal rotational axis. Afterward, vibrational and rotational energies are transformed into Cartesian coordinates and momenta following algorithms implemented in the VENUS package.[80,81] Random orientations in Euler angles between the (rigid body) Ar and the projectile (ion) are sampled in order to account for the random directions of the Ar formamide-M^{2+} collisions. Then the ion–projectile relative energy is set and possible impact parameters are considered. The impact parameter, b, is sampled between zero and b_{max} = 3.0 Å. Finally the collision is done at a given energy defined in the center-of-mass of the system composed by the ion and the projectile, E_{CM}. We considered three center-of-mass collision energies: 180, 230 and 280 kcal/mol, which match the experimental energy range of the available experiments.[82,83]

The trajectories were calculated using the VENUS[80,81] package coupled to Gaussian09.[84] The classical equations of motion were integrated using the velocity Verlet algorithm[85] with a time step of 0.2 fs that gives energy conservation for both reactive and nonreactive trajectories. The initial ion–Ar distance is 8.0 Åand the trajectories are stopped at a 100 Åion–Ar distance. This corresponds to a total integration time of about 2.5 ps *per* trajectory. A trajectory was also stopped if a reactive channel was identified. In that case

a criterion distance of 7.0 Åwas used to guarantee no interactions between fragments. 300 trajectories for each case were computed to correctly describe the process under study.

Dynamical reaction products. The main processes after the collision with Ar are: (i) conversion of collisional energy into internal energy of scattered ions; (ii) sequential activation of a bond and reactivity observed in the simulation time length (2.5ps); (iii) direct reaction after the collision (for example, an atom is knocked out). Fig. 41 summarizes the products obtained during the 2.5 ps simulations. The product ratios are also given in Table 11.

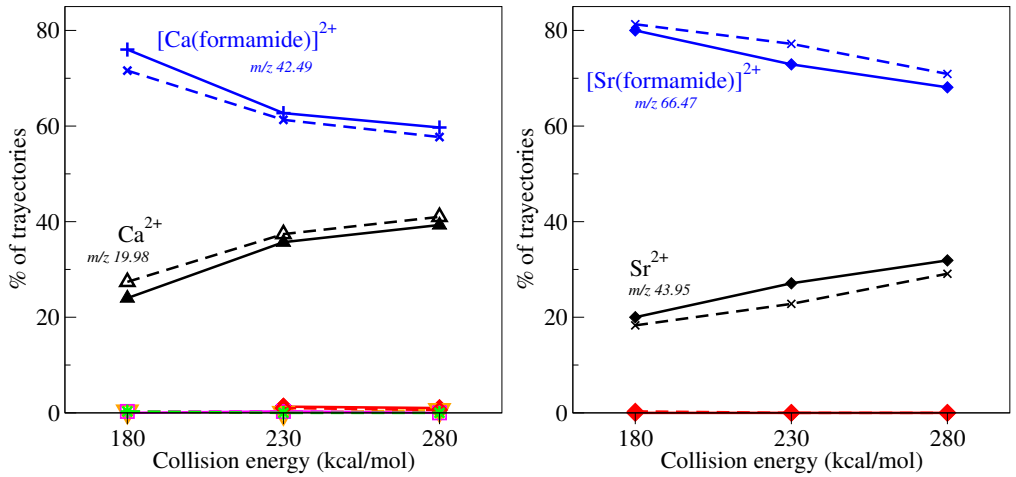


Figure 33: Percentages of trajectories for each channel as a function of the collision energy. For products observed in less than 2% see Table 11. Non-reactive trajectories (blue); product M^{2+} + formamide (black). The simulations were performed at two levels of theory: for M = Ca (left) G96LYP/6-31G(d) (solid lines) and BLYP/6-31G(d) (dashed lines). For M = Sr (right) G96LYP/6-31G(d) (solid lines) and G96LYP/6-31+G(d,p) (dashed lines).

The results do not significantly differ from one method to another. The first conspicuous fact is that the reactivity is not very high ($< 40\%$ of the trajectories react), and it is larger for calcium than for strontium. In the time length of our simulations (2.5 ps maximum) both kinds of reactive channels, neutral losses and Coulomb explosions, are observed. The former correspond to the loss of formamide and CO yielding as accompanying ions M^{2+} and $[Ca(NH3)]^{2+}$, respectively, and the latter to the fragmentation into

Table 11

Formamide-Ca^{2+}						
	G96LYP/6-31G(d)			BLYP/6-31G(d)		
Collision energy	180	230	280	180	230	280
Non reactive	76.0%	62.7%	59.7%	71.6%	61.3%	57.7%
Ca^{2+} + formamide	24.0%	35.7%	39.3%	27.4%	37.4%	41.0%
$[Ca(NH_2)]^+ + [HCO]^+$	—	1.3%	1.0%	—	1.0%	0.6%
$[Ca(NH_3)]^{2+}$ + CO	—	0.3%	—	0.3%	0.3%	—
int**10**	—	—	—	0.3%	—	—

Formamide-Sr^{2+}						
	G96LYP/6-31G(d)			G96LYP/6-31+G(d,p)		
Collision energy	180	230	280	180	230	280
Non reactive	80.0%	72.9%	68.1%	81.3%	77.2%	70.9%
Sr^{2+} + formamide	20.0%	27.1%	31.9%	18.3%	22.8%	29.1%
$[Sr(NH_2)]^+ + [HCO]^+$	—	—	0.3%	—	—	—

$[M(NH_2)]^+ + [HCO]^+$. Note that the loss of CO is only observed for M = Ca.

These results are in good agreement with CID experimental spectra (see Fig. 34).[82,83] Most of the trajectories did not react, which agrees with the fact that the peak for the parent ion is the most intense. Among the reactive ones, the vast majority undergo formamide neutral loss, the second most intense peak in the experiment. Similarly, neither the loss of CO nor the ions corresponding to the **G** Coulomb explosion are observed in the formamide-Sr^{2+} reactions, in agreement with our simulations. Indeed, no trajectories end in products **B** when M = Sr and only 0.3% of the trajectories follow the **G** Coulomb explosion pathway. It is worth noting however that some other products which appear in the experimental spectra did not show up in our chemical dynamics simulations, likely because the simulation time (2.5 ps) was too short. It should also be noticed that undeunderthe experimental conditions the ions might undergo multiple collisions while with direct dynamics trajectories we only model reactivity due to a single collision.

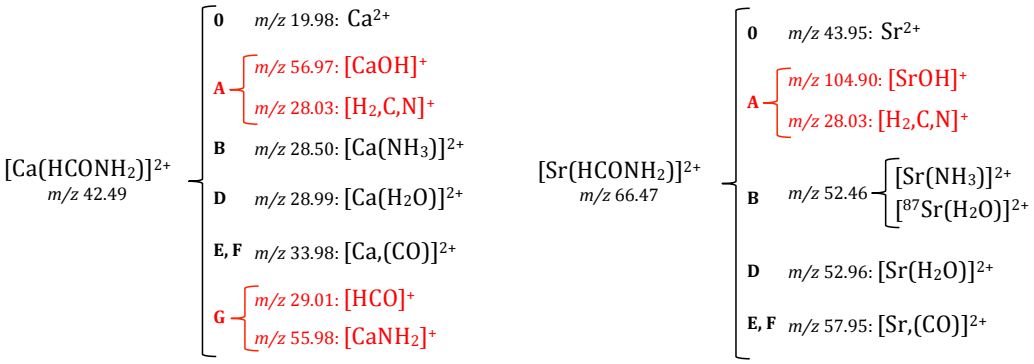


Figure 34: Masses observed in the experimental CID spectrum of formamide-M^{2+} together with the attributed structures. M = Ca (left), M = Sr (right). In black are neutral losses and in red Coulomb explosions. Data taken from ref.[82] and[83] respectively.

Effects of the collision energy in the reactivity. The agreement between our simulations and the experiments is also good as far as the effects of collision energy on the reactivity are concerned. Indeed, our theoretical results show that for all collision energies formamide neutral loss (O–M bond dissociation) clearly prevails over the other fragmentation pathways (which are observed for less than 1.5% of the trajectories), increasing almost linearly with the collision energy (see Fig. 41 and Table 11). This is consistent with a shattering mechanism occurring for this reactive channel (*vide infra*). For the **G** Coulomb explosion channel, giving rise to $[HCO]^+$ peak, the maximum number of trajectories following this path is observed at for 230 kcal mol^{-1} collision energy while it does not appear for the lowest collision energy used in the simulations (180 kcal mol^{-1}). Contrary, the **B** exit channel is not observed at the highest energy while it appears at low (BLYP) and middle energies (BLYP and G96LYP). Int**10** that appears for 180 kcal mol^{-1} collision energy (BLYP), is an intermediate structure between min**1** and product **B** that will eventually evolve to product **B**. The different effects of the collision energy on each reactive channel are better understood when considering the mechanisms for each pathway, as it will be discussed in the following section.

Reaction mechanisms An important feature of the dynamics simulations is that they provide an atomic-level description of the fragmentation mechanisms. There are three main ways in which the collision between the Ar

and the molecular ion takes place providing subsequent reactivity (Fig. 35): i) Ar hits the metal more or less perpendicularly to the M–O–C bond, ii) Ar strikes formamide molecule on the C=O bond and perpendicularly to the molecular plane or, iii) the collision takes place on the NH_2 group side and its rotation is thus activated.

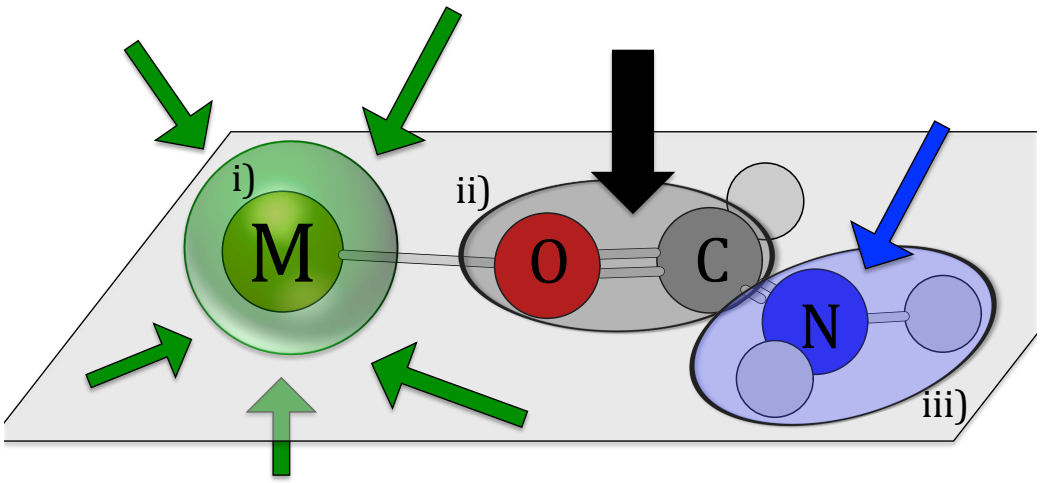


Figure 35: Different possibilities for the collision between Ar and the molecular ion providing subsequent dissociation, observed in our chemical dynamics simulations.

The vast majority of trajectories result in formamide neutral loss, regardless of the collision site. When Ar hits calcium perpendicular with respect to the M–O–C bond (i), the collision drives the metal away from formamide. This mechanism is similar to the "golf like" mechanism found by Spezia et al.**??** for CID of $[Ca(urea)]^{2+}$. When the collision is perpendicular to the C=O bond (ii), M^{2+} also detaches from formamide, breaking the M–O bond. Also, the M–O easily breaks for the collisions at the NH_2 group, when enough energy is transferred to the ion. Still, there is a difference due to the metal in the complex. In the case of Ca, more than one third of the reactive trajectories directly strip off the metal, whereas for Sr, which is more than twice heavier than Ca, the number of these trajectories is much smaller. A non-statistical mechanism, identified as "shattering",[86,87] occurs in surface-induced dissociation[88–90] (SID) where the projectile ion fragments as it collides with the surface. In a similar way, neutral loss of formamide in the formamide–M^{2+} systems studied here can be classified as a shattering mechanisms, since the M–O bond breaks due to the the Ar atom collision.

Therefore, the probability of breaking the M–O bond increases with the collision energy.

When the collision takes place at the NH_2 group (iii) also products **G** (Coulomb explosion), **B** (neutral loss), as well as the formation of int**10** are observed. The mechanism to arrive to **G** and **B** is very similar, with only very subtle differences between the two. In Fig. 36 the evolution of the natural charge with time for the two final fragments is plotted, as well as some representative distances, for one trajectory yielding **G** and another yielding **B** .

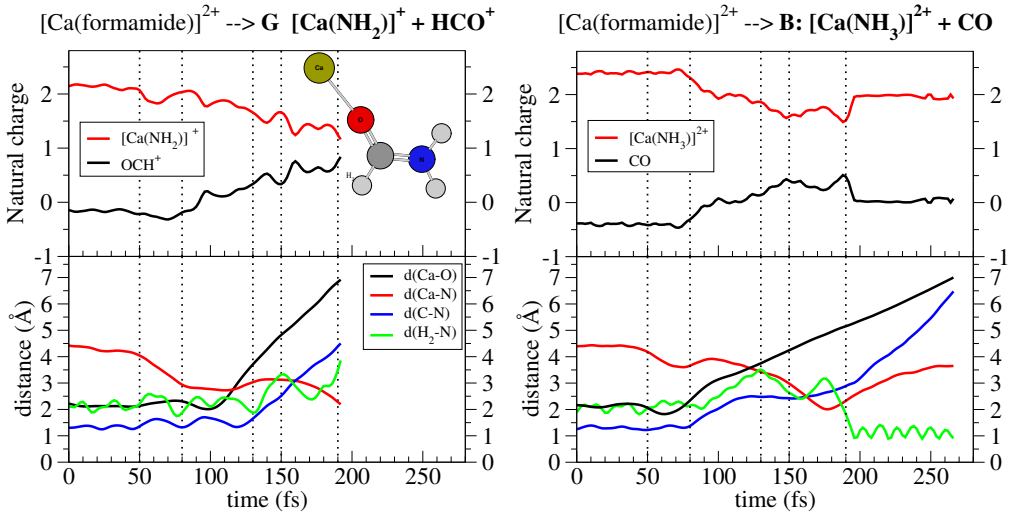


Figure 36: Evolution with time of natural charge (top panel) and distances (bottom panel) for a trajectory yielding **G**:$[Ca(NH_2)]^+$ + HCO^+ Coulomb explosion (left) and **B**: $[Ca(NH_3)]^{2+}$ + CO neutral loss (right).

In both processes the first 50 fs correspond to the approaching of Ar towards the molecular dication. After the collision in process **G** the first significant change is a decrease of the COCa angle, which brings the metal closer to the amino group, such that after 100 fs the Ca–N bond is practically formed and at that time both the Ca–O and the C–N bond start cleaving in an almost synchronous way. This leads to $[CaNH_2]^+$ and $[HCO]^+$ fragments, which repel each other as shown by the steep increase of both Ca–O and C–N distances. Coherently, the net charges of the two fragments start to change also significantly at ca. 100 fs. Conversely, for process **B** the first effect of the collision is the almost synchronous cleavage of Ca–O and C–N bonds,

so that after 100 fs Ca–O and Ca–N distances become very similar indicating that the metal is bridging between the HCO and the NH_2 groups. The separation of both subunits is not taking place until later. This allows the HCO group to reorient itself in order to favour a proton transfer towards the $CaNH_2$ moiety at around 200 fs, reflected by the sudden decrease of the N–H distance (light-green line). Thus, for low collision energies, there are more probabilities of obtaining **B** since the relative movement of the two subunits will be slow, facilitating the aforementioned proton transfer, whereas at high collision energies the Coulomb explosion should be greatly favoured.

Fig. 37 shows the simulation times (in fs) for chemical reactivity for all the trajectories. It is apparent that formamide neutral loss reaction covers a long time-span, from few femtoseconds up to ≈1.7 picoseconds, whereas the other reactions observed (**G** Coulomb explosion and CO neutral loss, **B**) are always fast, taking place in less than 1 ps. Looking into the details of each trajectory, we observe that the fast ($<$ 1 ps) formamide neutral loss reactions occur via a shattering mechanism, in which the energy is deposited into the M^{2+} + formamide relative motion leading to direct dissociation without IVR. However this is not the case for the remaining reactions, namely, "slow" formamide neutral loss, **G** Coulomb explosion and CO neutral loss, **B**. In these instances, there are some molecular rearrangements and/or energy distribution within the internal modes of the molecule. However, does a complete IVR occur? To answer this question we plotted in Fig. 38 the reaction times as a function of the internal energy for each trajectory yielding Ca^{2+} (left) and **G** Coulomb explosion (right), together with the half-life times ($t_{1/2}$) predicted with RRKM rate constants as a function of the internal energy of the molecule. The corresponding graphs for formamide neutral loss at the BLYP/6-31G(d) level, as well as for Sr^{2+} at G96LYP/6-31G(d) and G96LYP/6-31+G(d,p) are shown in Figs. A2.55 and A2.56 on Appendix 2, respectively.

Only in few cases for the Coulomb explosion reaction we have reaction times compatible with RRKM kinetics. However, in general, both sets of times differ significantly indicating that the trajectories which do not follow a shattering mechanism neither react through a full IVR mechanism, which is assumed in the RRKM theory. This suggests that the actual mechanism is in between these two limiting cases, *i.e.*, the energy is distributed within the internal degrees of freedom of the molecule, but the reaction takes place before a *complete* IVR could be achieved. Thus we called this mechanism "energy transfer", ET, as already proposed by Spezia et al.[91] These reactions are slower than *shattering* reactions, but still faster than predicted by

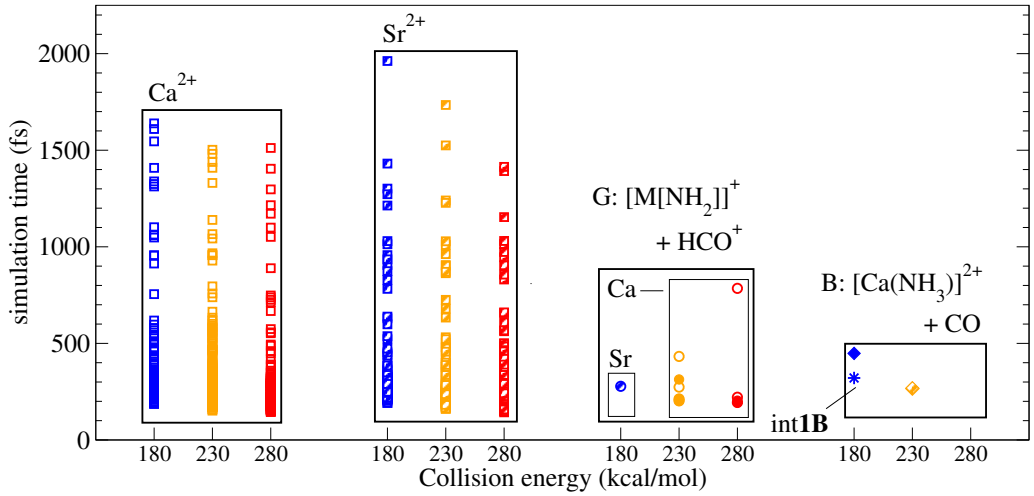


Figure 37: Simulation times (fs) for the reactive trajectories at the different collision energies: 180, 230 and 280 kcal mol^{-1}. Each square, circle or diamond stands for one trajectory yielding formamide neutral loss, **G** Coulomb explosion or CO neutral loss, respectively.

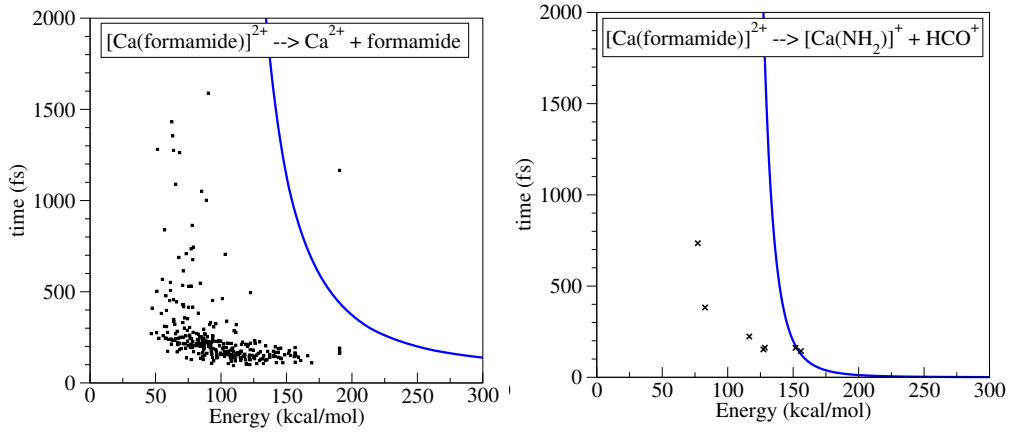


Figure 38: Reaction time *vs.* energy transfer obtained from chemical dynamics simulations (squares) and half-life times ($t_{1/2}$) predicted by RRKM (solid lines). Both were obtained using G96LYP/6-31G(d) level of theory. Results are shown for trajectories yielding formamide neutral loss (left) and **G** Coulomb explosion (right). M = Ca.

RRKM theory (complete IVR mechanism).

Non-reactive trajectories. Another piece of information that can be extracted from our chemical dynamics simulations is the amount of energy that has been transferred during the collision to the molecular ion, shown in Fig. S4 and S5 of the supporting information. Much more important, however, is to know how this energy is distributed after collision among the vibrational and rotational degrees of freedom for the min1 ion population that has not reacted in the 2.5 ps simulated by chemical dynamics. Fig. X shows the vibrational (a panels) and rotational (b panels) energy distributions for both Ca and Sr containing molecular ions, computed at the G96LYP/6-31G(d) level of theory. Similar results (Fig. S6 of the supporting information) are found with the other levels of theory.

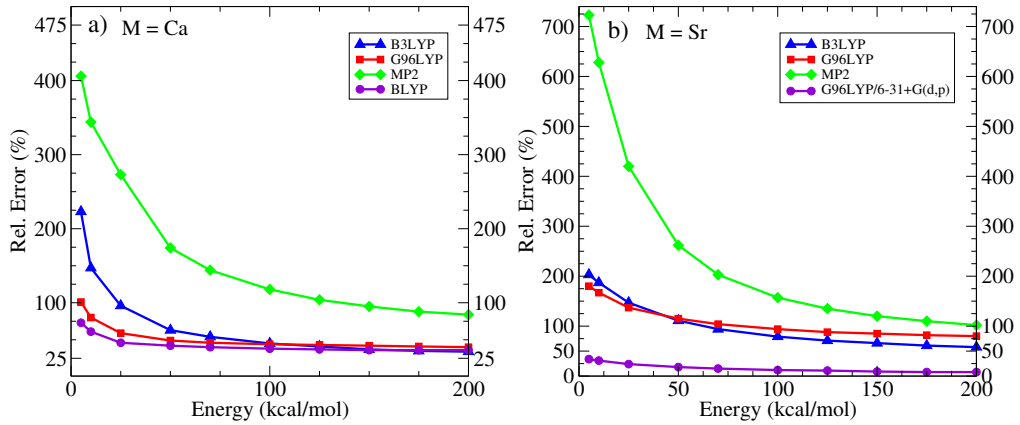


Figure 39

We should note that the vibrational activation for the molecules that did not react can be very important, reaching values up to 120 kcal/mol. One striking feature is the important amount of energy found in the rotational degrees of freedom after the collision with values up to 80 kcal/mol. Despite the fact that rotational activation is rarely considered in RRKM models applied to CID reactivity,26 it has been shown to be important in previous simulations of collisional activation of peptides,15a planar Al clusters,27 and molecular ions such as [Ca(urea)]2+ colliding with Ar,6e and protonated urea colliding with Ar25 and N2.15c In general, both vibrational and rotational excitations are slightly larger for [Ca(formamide)]2+ ions than for the Sr analogue. Due to the different masses, for Ca most of the ions have vibrational energies in the 10-20 kcal/mol interval while for Sr they are mainly in the 0-10 kcal/mol interval. In both cases the population decreases sharply as the energy increases. The same patterns are found for the rotational ex-

citation of both ions having most of the ions a rotational energy within the 0-10 kcal/mol range.

More importantly, for Ca there is a fraction of the non-reactive molecules with enough vibrational energy to cross the barriers between min1 and the five accessible TSs (dashed lines). Since the vibrational excitation is lower for Sr, only the lowest lying TSs can be reached. This suggests that G products are not observed in the CID experimental spectrum of [Sr(formamide)]2+ (in contrast to what is found for [Ca(formamide)]2+ ions) due to the lack of energy to reach the corresponding TSs.

Assuming that with enough time there will be extensive IVR, we can now couple the aforementioned chemical dynamics simulations with RRKM theory to build up a multi-scale approach that assesses longer time-scales. This coupling is simplified by the fact that vibrational excitation is independent from the rotational excitation (See the Supporting Information, Fig. S7 and S8).

The upper part of Fig. 9 shows the kinetic scheme for the reactions that can take place starting from min1, namely three isomerization reactions: min1 int2 (k12); min1 int5 (k15); and min1int10 (k110); and two fragmentation channels: formamide neutral loss (k01), and G Coulomb explosion (k1G). The bottom part of Fig. 9 shows the internal vibrational energy distributions, for the maximum collision energy used in experiments, obtained from the chemical dynamics simulations together with RRKM rate constants corresponding to the mentioned reactions. Similar graphs for the other collision energies considered can be found in the supporting information (Fig. S9). In a previous study8 we have shown that rotational energy has some effect only in k01, k1G, and k110 reaction constants when M = Ca, and k01, k1G, k110 and k15 when M = Sr. Hence, for these rate constants we have also plotted the curves for k(E) resulting from placing one third of the rotational energy in each of the three rotational axes with K quantum number treated adiabatically, i.e. energy exchange between vibrational and rotational modes is not allowed. The total rotational energy added is 27 and 18 kcal/mol for M = Ca and M = Sr, respectively (green curves).

For every collision energy all reactive channels starting from min1 are open in the case of [Ca(formamide)]2+ ion since the maximum values of internal vibrational energy are about 120 kcal/mol. However, only the first two channels are open (k01 and k12) when M = Sr, since the maximum values reached for the internal energy are about 70 kcal/mol. Only for Ecoll=280 kcal/mol a small fraction of [Sr(formamide)]2+ ions have energy enough to

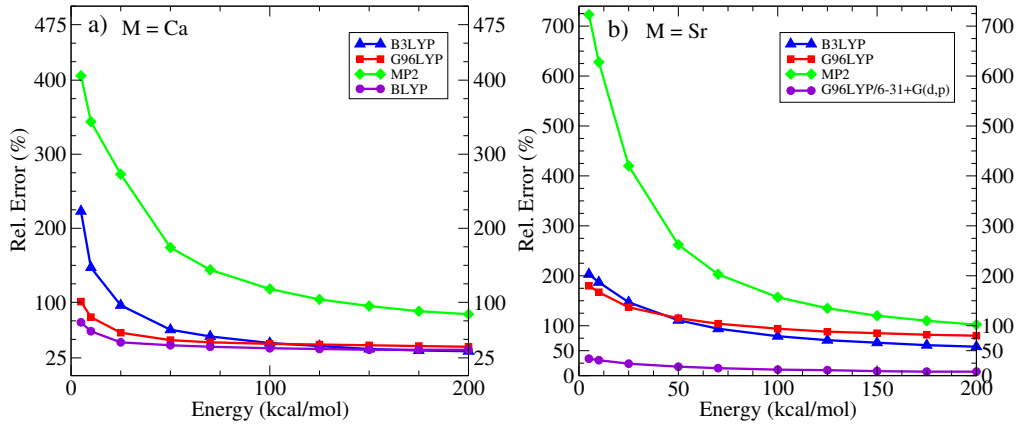


Figure 40

react through the k110 channel, that is open for energies greater than 75 kcal/mol. Thus, the difference between the two systems is mainly due to the difference in the internal vibrational energy distributions after collisional activation, given that the kinetic schemes for both metals are quite similar.

Knowing the internal vibrational energy distribution, it is possible to perform a kinetic analysis by means of the RRKM rate constants. In order to do so, we selected an upper value for the internal energy (vertical blue line in Fig. 9-b and -d) from the vibrational energy distributions corresponding to the maximum collision energy used in the experiments (Ecoll = 230 and 180 kcal/mol for Ca and Sr, respectively). For M = Ca the values for the various rate constants (blue arrows) in decreasing order are: k(84)01= 7.6109 s-1, k(84)12= 5.2108 s-1 , k(84) 110= 2.2107 s-1, and k(84)1G= 1.1105 s-1. The corresponding values for M = Sr are: k(65)01= 3.8109 s-1 and k(65)12= 1.7107 s-1, whereas there is not enough energy to follow the pathway associated with k15 when M = Ca and with k110, k1G and k15 for M = Sr (crossed out in the kinetic schemes of Fig. 9-a) and -c). For [Ca(formamide)]2+ at 84 kcal/mol, the fastest reaction, in the order of hundreds of ps, is the one leading to formamide neutral loss yielding Ca2+ followed by the ones leading to int2 (in the tens nanoseconds time scale) and to int10 (in the nanoseconds time scale). In the microsecond time regime, four orders of magnitude slower than formamide neutral loss, the G Coulomb explosion takes place. Hence, even the fastest reaction does not happen in the time-span covered by the direct dynamics simulations (¡ 2.5 ps) and consistently, the IVR mechanism is not observed during the simulations.

Finally, it is worth noting that rotational excitation may play also an important role in the dynamics of these systems. This effect has been illustrated in Fig. 9, when 27 (Ca) and 18 (Sr) kcal/mol of rotational energy is added to k01, k1G and k110 reactions, the only ones in which rotational excitation can be important.8 It is apparent that for both metal ions all the rate constants decrease and the curves appear shifted to higher energies, this shifting being different depending on the process. As a consequence, for M = Ca, the Coulomb explosion G becomes now slower than the isomerization to yield int10, whereas k01 and k12 become competitive at E = 84 kcal/mol. For M = Sr the effects are less dramatic, because although k01 is also shifted to higher energies, it does not cross k12, and the other reactions do not take place.

This multi-time-scale approach can also be applied to investigate the evolution of the intermediates that can be reached from min1. This will be illustrated, as a suitable example for int2. The upper part of Fig. 10 shows the corresponding kinetic schemes for both metal ions. In the bottom part are represented the RRKM rate constants for the matching reactions.

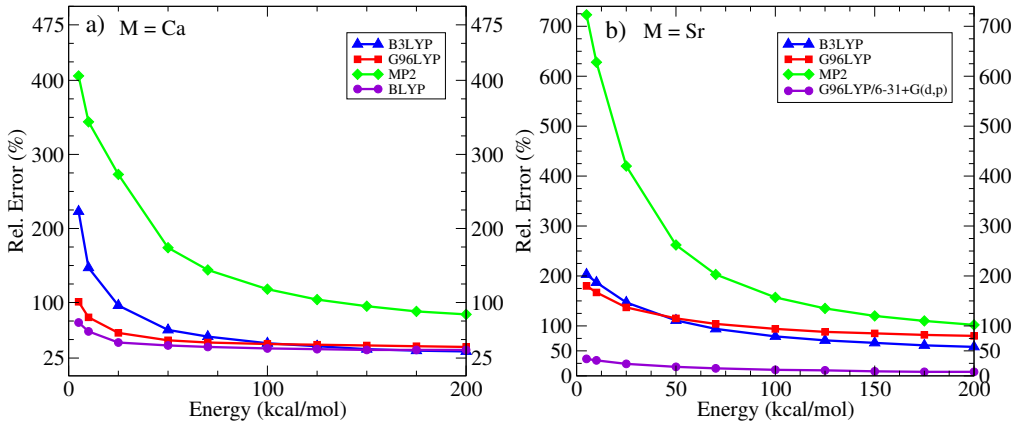


Figure 41

From int2 there are three possible options: going back to the origin, min1; fragmentation of int2 via Coulomb explosion into products A, and an isomerization to int3 that would eventually loose ammonia to provide product D.9 For the sake of consistency, the curves were shifted to the left of a quantity corresponding to the relative energy of each intermediate with respect to min1. The first conspicuous observation is that for both Ca and Sr the backward reaction to min1 (k21) is faster than the forward one (k12); but

reactions towards int3 (k23) and products A are even faster. There are, however, some subtle differences between the two metals as far as the reaction to A products is concerned. Whereas for M = Ca this reaction is much faster than the one going back to min1 and that yielding int3, for M = Sr it is slower than k23 and competes with k21. Thus, for M = Ca, int2 would rapidly evolve to products A and a very small fraction would isomerize to int3, whereas for M = Sr most of the flux will follow the path leading to int3. In both cases this intermediate will eventually evolve to product D: [M(H2O)]2+ (M = Ca, Sr). Similar analyses have been performed for the intermediate int10 that eventually yields product B: [Ca(NH3)]2+ (see the Fig. S10 in the supporting information).

Summarizing, M2+ ions are obtained from both fast dynamical shattering and slow IVR statistical mechanisms. Following statistical mechanisms, the two systems will also form A Coulomb explosion products: [MOH]+ + [HCNH]+ and D neutral loss product: [M(H2O)]2+. However, formation of D will be dominant with respect to the formation of A for M = Sr, whilst the opposite is true for the M = Ca, in agreement with the experimental evidence. For M = Ca, in which the vibrational excitation of the ions is higher, two more reaction channels are open, B: [Ca(NH3)]2+ and G: [Ca(NH2)]+ + [HCO]+, which are however not accessible when M = Sr, again in agreement with the experimental observations. Therefore, the un-assigned peak at m/z 52.46 in the [formamide-Sr]2+ mass spectra should correspond to [87Sr(H2O)]2+ rather than to [Sr(NH3)]2+, contrary to what we have previously suggested based on static calculations only.

Concluding remarks. The fragmentation of [formamide-M]2+ complexes follows both statistical and dynamically driven mechanisms. Indeed, almost half of the trajectories reacts in a short time scale following a dynamical mechanism. The time to form the different products varies in a wide range, from few femtoseconds to nanoseconds. In the short-time scale (¡2.5ps) we observed mainly dynamically driven reaction (shattering mechanisms), as formamide neutral loss. Coupling the internal vibrational energy distributions obtained from chemical dynamics simulations with RRKM rate constants we could explore phenomena occurring at longer time-scales. Thus, we could also account for slow reactions occurring via IVR mechanisms and leading to a wider range of products.

The use of direct dynamical simulations allowed us to explain the presence

of the M2+ peak, otherwise impossible to explain based on PES analyses only (i.e. corresponding to the most endothermic exit channel). Furthermore, the dissimilarities between the two molecular ions such as the larger reactivity of the [Ca(formamide)]2+ system or the absence of G Coulomb explosion in [Sr(formamide)]2+ fragmentations, that cannot be explained neither from the PES topology nor using RRKM theory, can also be accounted for only when using the dynamical formalism. Therefore, by combining the three approaches; CID dynamics simulations, static (PES) and statistical (RRKM) analysis, we could explore a wide range of reaction time-scales and ultimately account for all the products observed in the CID experimental spectra of [Ca(formamide)]2+ and [Sr(formamide)]2+ doubly-charged cations, as well as explain the differences observed between these ions.

4 Conclusiones

La fragmentacin de los iones formamida-M^{2+} siendo M = Ca, Sr, sigue diferentes mecanismos estadsticos y dinmicos (no estadsticos). De echo, aproximadamente la mitad de las trayectorias reacciona en una escala de tiempor corta siguiendo mecanismos dinmicos. Los tiempos observados para la formacin de los distintos fragmentos/productos vara en una amplia escala de tiempo que abarca desde femtosegundos hasta nanosegundos. Acomplando las distribuciones de energa vibracional y rotacional para los iones que no reaccionan durante el tiempo simulado con las trayectorias de dinmica con las constantes de reaccin microcannicas obtenidas usando la teora RRKM, hemos sido capaces de explorar escalas de tiempo laras (> fs). De este modo hemos sido capaces de describir la reacciones “lentas” que tiene lugar por mecanismos estadsticos que implican una completa redistribucin de la energa dentro de la molecula despus de que la colisin tenga lugar y por tanto, hemos sido capaces de explicar la aparicin de una mayor nmero de fragmentos/productos. Mediante la combinacin de las distintas tcnicas tambin hemos sido capaces de explicar las diferencias observadas entre los dos metales, Ca y Sr, cosa imposible utilizando aisladamente cualquiera de las tcnicas mencionadas a lo largo de este manuscrito.

METHODOLOGY

1 Theoretical methods

Quantum Chemistry aims to interpret and predict the electronic structure and reactivity of chemical systems, using the principles of Quantum Mechanics. With this goal, since Schrödinger solved the hydrogen atom equation in 1926, several approaches with increasing levels of complexity have been developped. At the same time, more and more efficient computational programs have been developed, using at each moment the state of the art informatics technology to deal more accurately with systems of ever increasing complexity and size.[?]

To overcome the impossibility of exactly solving the Schrödinger equation for many electron systems, in 1930 Hartree and Fock proposed the Self-Consistent-Field (SCF) method, the origin of the current *ab initio* calculations. However, the SCF calculations were not to become truly operative until the introduction of the Linear Combination of Atomic Orbitals (LCAO) approximation. The introduction of a basis set transformed the numerical problems of solving integro-differential equations into the well-known resolution of linear algebra equations.

SCF is only a first approach to most chemical systems, because it is generally insufficient for correctly describing systems with unpaired electrons or breaking bonds. Another of its drawbacks is that it considers electrons to be independent particles that move in an average field, thus the electrons movements are not correlated within this approach. In order to address this problem different methods were the electron correaltion is included have been developed. These methods start with an orbital set usually obtained by a self-consistent method. Since they include the electron correlation *a posteriori*, they are often referred to as *post* Hartree-Fock methods. They belong to three main categories: Configuration Interaction methods (CI), Coupled Cluster (CC) methods and Many body Perturbation Theory methods (PT).

In 1964, Hohenberg and Kohn showed that the ground state energy depends only on the electron density and thus they established the basis of the Density Functional Theory (DFT). A year later, the Kohn-Sham equations enabled the application of this method to molecular and solid state systems. In recent years, DFT applications have grown spectacularly because the good results that can be obtained at low cost. Its main drawbacks are due to the fact that it is a single configuration methodology and there is not a systematic way to improve its performance.

1.1 Wave function-based methods.

The existence of matter waves suggests the existence of a wave equation describing them. Such a wave equation was first proposed by the Austrian physicist Erwin Schrödinger in 1926.[?] According to this approach, for any system it exists a function Ψ, namely wavefunction, that contains all the information about the system. A particular solution for this equation is (4). It is known as the *time-independent Schrödinger* equation and it represents *stationary states* of the system.

$$\hat{H}\Psi = E\Psi \tag{4}$$

where $\hat{H}$ is called the *hamiltonian operator*, $\hat{H} = -\frac{\hbar^2}{2m}\nabla^2 + \hat{V}$. Due to the presence of the ∇^2 operator, equation (4) can be solved only if it can be separated on its different variables. In the $\hat{V}$ operator there are terms involving the coordinates of more than one particle at the same time and consequently the Schrödinger equation cannot be solved in an analytical way and we need to consider different approximations in order to obtain approximate solutions.

Born-Oppenheimer approximation. Born and Oppenheimer made the first of these approximations in 1927.[?] It is based on the large disparity in mass between nuclei and electrons (on the order of 2000 to 1 or more). Since the nuclei are much more massive than the electrons, they tend to move considerably more slowly than do the electrons. We assume that the electrons readjust themselves very rapidly to any configurational change of the nuclei. In the Born-Oppenheimer approximation we assume that both movements are uncoupled. Thus, first we fixed the nuclei positions and consider the electrons movements around them. We obtain an energy value for the electrons in the fixed nuclei configuration. This energy together with the

internuclear repulsion makes up the potential energy to which the nuclei are subject. With this potential and the kinetic energy for the nuclei we can solve the nuclear part of the equation and thus calculate the total energy and wavefunction for the system.

The orbital approximation. As alluded to previously, for an N-electron system the Schrödinger equation cannot be solved exactly and we need to consider approximate wave functions. One of the simplest models used to describe many-electron systems is known as the *orbital approximation.*[?] In this approximation the hamiltonian is assumed to be of the *spin-free* form, and the many-electron wavefunction is assumed to have the general form of a determinantal product of orbitals:

$$\Psi(1,2,...,N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \phi_1(1) & \overline{\phi_1}(1) & .... & \overline{\phi_N}(1) \\ \phi_1(2) & \overline{\phi_1}(2) & .... & \overline{\phi_N}(2) \\ .... & .... & .... & .... \\ \phi_1(N) & \overline{\phi_1}(N) & .... & \overline{\phi_N}(N) \end{vmatrix} \tag{5}$$

A wavefunction of this form is called a *Slater determinant.*[?] The quantity $\phi_i(\mu)$ is called a *spin orbital* (product of the spatial part and the spin function, see eq.(6)) and μ is used to designated each electron. A spin orbital is a one-electron distribution function involving the full coordinates (three spatial and one spin) of an electron.

$$\chi_1(1) = \phi_1(1)\alpha(1) or \chi_1(1) = \phi_1(1)\beta(1) = \overline{\chi}_1(1) \tag{6}$$

The Slater determinant is the simplest form for an antisymmetric wavefunction. Since a determinant will change sign when any two of its rows are interchanged, such a wavefunction will incorporate the antisymmetry required whenever two fermions have their full coordinates interchanged. It also fulfils the restriction known as the Pauli exclusion principle: two electrons cannot have the same full set of quantum numbers, because if there are two identical columns (or rows) the determinant is equal to zero.

Linear expansion of the wavefunction. If we consider a complete basis set $\{\chi_i(x)\}$, any one-electron function $\Phi(x)$ can be exactly expanded using this basis set

$$\Phi(x) = \sum_i a_i \chi_i(x) \tag{7}$$

where a_i are the coefficients for the linear combination. Let us suppose now a function of two electrons, in principle with a fixed second variable, x_2. This function can be expanded as:

$$\Phi(x_1, x_2) = \sum_i a_i(x_2)\chi_i(x_1) \tag{8}$$

Since the expansion coefficients are also functions of a single variable, they can be expressed as:

$$a_i(x_2) = \sum_j b_{ij}\chi_j(x_2) \tag{9}$$

and

$$\Phi(x_1, x_2) = \sum_i \sum_j b_{ij}\chi_j(x_2)\chi_i(x_1) \tag{10}$$

To satisfy the Antisymmetry principle, $b_{ij} = -b_{ij}$ and them

$$\begin{aligned} \Phi(x_1, x_2) &= \sum_i \sum_{j>i} b_{ij}[\chi_i(x_1)\chi_j(x_2) - \chi_i(x_2)\chi_j(x_1)] \\ &= \sum_{j>i} b_{ij}\frac{1}{\sqrt{2}} \begin{vmatrix} \chi_i(x_1) & \chi_j(x_1) \\ \chi_i(x_2) & \chi_j(x_2) \end{vmatrix} \end{aligned} \tag{11}$$

Thus, an N-electron wavefunction can exactly be expanded in a linear combination of all the determinants obtained from N one-electron functions.[?] As far the one-electron basis-set is complete, the expansion is exact. The problem now is to choose this basis set, because the wavefunction and the energy depend on it.

1.1.1 The Hartree-Fock Self-Consistent-Field Method

Within the Hartree-Fock method framework the variational theorem is used to determine the electronic energy of a molecular system using as trial function a Slater determinant composed of orthonormal molecular spin-orbitals.

Hartree introduced the first procedure for this kind of calculations in 1928[?] and is called the *Hartree SCF method*. Hartree arrived at the SCF

procedure by intuitive physical arguments. The proof that Hartree's procedure gives the best possible variation function was given by Slater[?] and by Fock[?] in 1930, given rise to the actual name of the method: Hartree-Fock approximation. Within this approximation, the expansion of the wavefunction Φ in a linear combination of determinants (see eq. (11)) is truncated to a single term Φ_0, whose elements are molecular orbitals.

Variational theorem. The importance of this theorem lies on the fact the it gives a systematic way to improve the energy of a system obtained by means of a trial function. It says that the energy associated to a normalized wavefunction Φ_0 as an associated energy W which is an upper limit to the ground state energy of the system E.

$$W = \langle \Psi \left| \hat{H} \right| \Psi \rangle \geq E \tag{12}$$

So now we have a systematic way of improving the trial wavefunction used as a solution of the Schrödinger equation. It is worth to note that a wavefunction which gives the best energy estimate does not necessarily do the same for other properties of the system.[?]

Energy minimization The general problem is to seek those spin orbitals which make the total energy $\langle \Psi \left| \hat{H} \right| \Psi \rangle$ a minimum. There are no restrictions on the spin orbitals other than that they lead to a well-behaved wavefunction. However there are other imposed restrictions: The determinant is to be of closed-shell form, with each spin orbital expressed as a product of a spatial orbital ϕ and a spin function (α or β). An additional restriction, introduced to facilitate numerical computations, the spatial orbitals are chosen to be orthonormal, namely

$$\langle \phi_i | \phi_j \rangle = \delta_{ij} \tag{13}$$

The total energy of a system having a wavefunction of the form 5 and consisting on N doubly occupied spatial orbitals is given by:

$$E = 2 \sum_{i=1}^{N} \epsilon_i^{(0)} + \sum_{i,j}^{N} (2J_{ij} - K_{ij}) \tag{14}$$

We now wish to find the best possible orbitals (subject to the restrictions mentioned) to use in a wavefunction of single-determinantal form; "best possible" means "leading to the lowest energy". For this the lagrangian multipliers technique will be used.

The treatment is greatly simplified in a notational sense if we define coulomb and exchange operators as follows:

$$J_i(\mu)\phi_j(\mu) = \langle \phi_i(\nu) \left| \frac{1}{r_{\mu\nu}} \right| \phi_i(\nu) \rangle \phi_j(\mu) \tag{15}$$

$$K_i(\mu)\phi_j(\mu) = \langle \phi_i(\nu) \left| \frac{1}{r_{\mu\nu}} \right| \phi_j(\nu) \rangle \phi_i(\mu) \tag{16}$$

With the above definitions of the coulomb and exchange operators, the coulomb and exchange integrals (both representing electron repulsions) occurring in the expression given by eq. (14) can be written:

$$J_{ij} = \langle \phi_i(\nu) \left| J_i(\nu) \right| \phi_i(\nu) \rangle = \langle \phi_j(\mu) \left| J_i(\mu) \right| \phi_j(\mu) \rangle \tag{17}$$

$$K_{ij} = \langle \phi_i(\nu) \left| K_i(\nu) \right| \phi_i(\nu) \rangle = \langle \phi_j(\mu) \left| K_i(\mu) \right| \phi_j(\mu) \rangle \tag{18}$$

It is seen that the coulomb operator is just the operator for the potential energy, which would arise from an electron distribution $|\phi_i|^2$. Such operators represent the effective potentials for an electron moving in the repulsive field of other electrons. The exchange operator, on the other hand, has no classical analogue, since it arises from the non-classical antisymmetry principle.

After some mathematical treatment that we will not discus here we obtain the *Hartree-Fock equations*:

$$\left[\hat{h}_i^{core}(1) + \sum_j^{n/2} \left(2\hat{J}_j(1) \right) \right] \phi_i(1) = \sum_j^{n/2} \epsilon_{ij} \phi_j(1) \tag{19}$$

where (1) emphasizes that the orbital or operator depend on the coordinates of one single electron.

The Hartree-Fock equations can be written in a matrix form

$$\hat{F}\Psi = E\Psi \tag{20}$$

$$\tag{21}$$

where $\hat{F}$ is the Hartree-Fock operator defined by

$$\hat{F} = h_\nu + \sum_j^N (2J_j - K_j) \tag{22}$$

$$\tag{23}$$

It is important to note that the sole function of the Hartree-Fock operator is to generate the orbitals to be used in the wavefunction given by (5). Once these orbitals are obtained, they are used in 14 to calculate the energy of the system.

Also, it is worth to mention that **E** in eq.(21) is not a diagonal matrix. However, a unitary transformation of these orbital will leave the total wavefunction invariant, we can find a unitary transformation which diagonalizes the matrix **E** and transform eq.(21) in a pseudo-eigenvalue equation

$$\hat{F}\Psi_i = \epsilon_i \Psi_i \tag{24}$$

the $\{\epsilon_i\}$ are called the *Hartree-Fock eigenvalues* and are the energies associated in a specific way with the Hartree-Fock orbitals.

Since the operators $J_i(\nu)$ and $K_i(\nu)$ appearing in the Hartree-Fock operator depend on the orbitals generated by this operator, it is not possible to solve the Hartree-Fock equations in the same manner as a true eigenvalue equations are solved, hence the reference to these as *pseudo-eigenvalue* equations. One of the oldest procedures used to solve these equations is the *self-consistent field* (SCF) method. The basic procedure is to choose some beginning set of orbitals $\{\phi_i^{(0)}\}$, which are used to construct an initial approximation of the coulombic and exchange operators found in $\hat{F}$. A first approximation to the optimized orbitals is then obtained by solving

$$\hat{F}^{(0)}\Phi_i^{(1)} = \epsilon_i^{(1)}\Phi_i^{(1)} \tag{25}$$

The new orbitals $\{phi_i^{(1)}\}$ are now used to redefine the Hartree-Fock operator, and a second, further improved set of orbitals is obtained by solution of

$$\hat{F}^{(1)}\Phi_i^{(2)} = \epsilon_i^{(2)}\Phi_i^{(2)} \tag{26}$$

At some point the functions obtained in an iteration become virtually the same as those of the subsequent iteration (to within some specified tolerance limit), and we say a *self-consistent field* has been reached. It is this final set of SCF orbitals, which constitute the optimized orbitals, used to calculate the energy via equation (14).

The physical significance of the eigenvalues of the Hartree-Fock operator $\hat{F}$ is made clear by a theorem first proved by the Dutch physicist T.C. Koopmans.[?] Koopmans showed that the optimized orbitals used to construct the

single-determinantal wavefunction of an atom X and its two ions X^+ and X^- were the same to within same order. Thus, the energies of the ions are related to that of the neutral atoms by the simple relationships

$$E(X^+) = E(X) - \epsilon_k \qquad k = 1, 2, ..., N \tag{27}$$

$$E(X^-) = E(X) + \epsilon_m \qquad m = N + 1, N + 2, ... \tag{28}$$

where k refers to an orbital used to describe the ground state of the neutral atom and m refers to what is called a *virtual* orbital. The expression given by (28) thus imply that the energies of the occupied orbitals should be approximations to the negatives of various ionization energies of the neutral atom. Thus, if ϕ_k is the highest (in terms of energy) occupied orbital of the ground state $-\epsilon_k$ is an approximation of the first ionization energy of the atom. Similarly, the energies of the virtual orbitals are approximations of the electron affinities of the atom.

Such solutions generally involve numerical integration techniques, and the orbitals themselves are not expressible in the usual analytical form but rather in terms of tabulated numerical values over a grid of spatial positions. A great deal of computation is required to perform a Hartree-Fock SCF calculation for many-electron atom. Hartree did several SCF calculations in the 1930s, when electronic computers were not in existence. Fortunately, Hartree's father, a retired engineer, enjoyed numerical computation as a hobby and helped his son.[?] In 1951, Roothaan[?] and Hall[?] proved, independently, that this problem can be solved if the molecular orbitals are expressed as a lineal combination of basis set functions.

$$\Psi_i = \sum_s c_{is}\phi_s \tag{29}$$

or in a matrix way

$$\Psi = C\phi \tag{30}$$

where C is the matrix for the coefficients of the expansion and $\{\phi\}$ the set of basis functions. The representation of the molecular orbitals as in (29) is a purely chemical point of view: as a molecule is composed by atoms, a molecular orbital can be composed by a set of atomic orbitals. That is the reason why this approximation is known as *Linear Combination of Atomic Orbitals* (LCOA).

The transcription of the Hartree-Fock equations when an atomic orbital basis set has been defined to expand the spatial part of each spin orbital

gives rise to the Roothann-Hall equations. Before applying this approximation, the decision of adopting or not the spin restriction has to be taken. For systems with an even number of electrons in their ground state, the closed-shell systems, the spin restriction is the most appropriate form, and the method is known as *Restricted Hartree-Fock* (RHF). For open-shell systems, we can choose one solution or the other. If we maintain the restriction, the method is known as *Open-Shell Restricted Hartree-Fock*(OSRHF), and otherwise, *Unrestricted Hartree-Fock* (UHF).

The equations for a closed-shell systems are known as the Roothaan-Hall equations.

$$FC = SCE \tag{31}$$

in the compact matrix form. **F** is the matrix representation of the Fock operator in the atomic orbital basis, **c** the vector of coefficients of the expansion for the corresponding orbital ϕ_i, **C** the matrix of columns $\{c_a\}$ and **E** the diagonal matrix of the energies $\{\epsilon_i\}$. On the other hand, if spin restriction is not taken into account, the UHF method leads to two series of equations analogous to (31), the Pople-Nesbet equations:

$$F^\alpha C^\alpha = SC^\alpha E^\alpha F^\beta C^\beta = SC^\beta E^\beta \tag{32}$$

These two sets of equations are not independent since the two Fock matrices depend on the total density matrix, **P**, and they must be solved simultaneously.

Hartree-Fock method limitations Energies calculated by the Hartree-Fock method are typically in error by about 0.5% for light atoms. On an absolute basis this is not much, but for the chemist is too large. For example, the total energy of the carbon atom is about -1000 eV, and 0.5% of this is 5 eV. Chemical single-bond energies run about 5 eV. Calculating a bond energy by taking the difference between Hartree-Fock molecular and atomic energies, which are in error by several electronvolts for light atoms, is an unreliable procedure, consequently, a way to improve Hartree-Fock wave functions and energies must been developed.

A Hartree-Fock SCF wave function takes into account the interactions between electrons only in an average way, but the instantaneous interactions between electrons must be considered as well. Since, the electrons repel each other, they tend to keep out of each other's way. Thus, the motions of electrons are correlated with each other, and we speak of *electron correlation.*

Different methods have been developed in order to include the instantaneous electron correlation into the wave function.

Actually, a Hartree-Fock wave function does have some instantaneous electron correlation because it satisfies the antisymmetry requirement. Therefore, for a Hartree-Fock function there is a little probability of finding electrons of the same spin in the same region of space, so it has some correlation of the motions of electrons with the same spin. This region in which the probability of finding another electron with the same spin is small is usually referred to as *Fermi hole*. We can also speak of a *Coulomb hole* surrounding each electron in an atom. This is a region in which the probability of finding another electron is small.

We can define a quantity called the *electron correlation energy*, which is a measure of the ability (or inability) of the Hartree-Fock wavefunction to provide an accurate description of the electronic structure of an atom (or molecule). After Löwdin, we define the electron correlation energy as follows: The correlation energy of a certain state with respect to a specified Hamiltonian is the difference between the exact eigenvalue of the Hamiltonian and its expectation value in the Hartree-Fock approximation for the state under consideration.[?] The correlation energy defined in this way depends on the hamiltonian and may be expressed mathematically as

$$E_{corr} = \langle \hat{H} \rangle (exact) - \langle \hat{H} \rangle (Hartree - Fock) \tag{33}$$

There are several methods that introduce the electronic correlation. The conventional ones start with the Hartree-Fock wavefunction and consequently they are often referred to as *post* Hartree-Fock methods. Within these methods the most well known are: Configuration Interaction methods (CI), Coupled Cluster (CC) methods and the perturbational method of Møller-Plesset (MP).[?]

There are other no conventional methods that include the electronic correlation in an alternative way. These are the methods based on the *Density Functional Theory* (DFT) that will be discussed later.

1.2 Atomic basis sets

The LCAO formulation, which is used almost universally in quantum chemistry methods, requires the definition of a set of functions to expand the

spatial part of spin orbitals according to equation (29). Thus, most molecular quantum-mechanical methods, whether SCF, CI, perturbation theory, coupled cluster or density functional, begin the calculation with choice of a set of basis functions. This set is normally known as the *atomic basis set*. In fact, any mathematical function can be exactly represented as a linear combination of basis functions, if the set is complete, which unfortunately involves an infinite number of functions. In practice, then, we have to do with an approximate representation of our orbitals since we necessarily use a finite basis. The use of an adequate basis set is an essential requirement for success of the calculation. The two most commonly used orbitals are described in this section.

Slater-type orbitals (STO) At the dawn of quantum chemistry, Slater proposed using functions that, although simpler than hydrogen like orbitals, preserved the same exponential dependence on the electron-nucleus distance, r. These functions, known as Slater-type orbitals (STO)[?] have the general form:

$$\phi_{\zeta,n,l,m}(r,\theta,\varphi) = Nr^{n-1}e^{\zeta r}Y_{l,m}(\theta,\varphi) \tag{34}$$

where N, is a normalization constant and $Y_{l,m}$ the spherical harmonics.

Although the exponential dependence on r guarantees a good description of the function maximum (*cusp*) on the nucleus (r$\rightarrow$0) and a rapid convergence with the number of functions, the calculation of tri-and tetra centric integrals is not analytical, so the use of STO has been restricted almost exclusively to the treatment of atoms and diatomic molecules.

Gaussian-type orbitals (GTO) To speed up molecular integral evaluation, Boys proposed in 1950[?] the use of Gaussian-type functions instead of STOs for the atomic orbitals in an LCAO wave function. A Cartesian Gaussian centred on atom b is defined as

$$g_{ijk} = Nx_b^i y_b^j z_b^k e^{-\alpha r_b^2} \tag{35}$$

where i, j and k are nonnegative integers, α is a positive orbital exponent, x_b, y_b and z_b are Cartesian coordinates with the origin at the nucleus b, and r_b is the distance to the nucleus b . The sum i + j + k determines the type of orbital (for example, $l_x+l_y+l_z = 0$ is an s-type orbital, whereas $l_x+l_y+l_z = 1$ is a p-type orbital).In general the GTOs have the basic deficiency that they

do a very poor job of representing the electron probability both near the nucleus and far away from it. However, this deficiency may be overcome by using a large number of GTOs. For example, a linear combination of several GTOs will essentially replace a single STO.[?]
A number of unconventional basis sets have been used by various investigators over the course of time, but none has enjoyed the overall utility of STOs. Some basis sets may appear to surpass STOs when employed for certain types of simple calculations, but these quickly lose their advantages when used in more complicated wavefunctions.[?]

1.2.1 Types of basis sets

Minimal basis set Let us discuss some of the terminology used to describe STO basis sets. A minimal basis set consists of one STO for each inner-shell and valence shell atomic orbital of each atom. For example, for C_2H_2 a minimal basis set consists of 1s, 2s, $2p_x$, $2p_y$ and $2p_z$ atomic orbitals on each carbon and a 1s STO on each hydrogen. There are five STOs on each C atom and one on each H, for a total of 12 basis functions. This set contains two s-type STOs and one set of p-type STOs on each carbon and one s-type STO on each hydrogen. Such a set is denoted by (2s1p) for the carbon functions and (1s) for the hydrogen functions, a notation, which is further abbreviated to (2s1p/1s).

An example of this kind of basis set is the STO-NG basis set generated by Pople and co-workers as an effective way of conjugating the advantages of STO at short and long distances and the advantage of using Gaussian functions. Thus, all the integrals are analytical. The idea is very simple and consists of using least squares to fit an expansion of N Gaussian functions to a particular STO:

$$\phi_{STO-NG} = \sum_{l}^{N} a_i \chi_i \tag{36}$$

where χ_i is a Gaussian function and the coefficients of the expansion, a_i, are obtained as mentioned above by least squares fitting to a particular STO. Although the expansion can include as many terms as are desired it is easily verified that the description of the corresponding STO for $N > 3$ improves very little in comparison with the increase in computational cost. So, the most widely used expansion is STO-3G. This means that, when the basis set of functions used $\{\phi_i\}$, what is in fact being used for each function is an

expansion such as the one given in equation (36) in which the coefficients are predetermined. Therefore, these types of expansions are usually called *contracted Gaussian functions* and the functions χ_i, that are used in the contraction are called *primitive Gaussian functions*. A STO-3G, for instance, is a contracted Gaussian made up of three primitives.

The advantage of the minimal basis sets is that, due to the fact that they are composed by few functions, they are very fast to evaluate, so they are quite useful to have a qualitative idea of the problem. However, they are not very flexible and, therefore, if we used a correlated method the amount of electronic correlation that we recuperate is lower than when we use a bigger basis set. Consequently, they are not appropriate to obtain very accurate quantitative results.

1.2.2 N-zeta and Split-Valence basis sets

When a bond is formed between two different atoms, the charge is not distributed equitatively. Thus, leading to the bond polarization and a different distribution of the atom's charge within the molecule. For example, the atom that has a larger partial positive charge would have a more contracted electronic cloud while for the other one, the electronic cloud would be more disperse.

Another common feature of the charge distribution in a bond is the anisotropy. For instance, in a multiple bond, the π orbitals are more diffused than in the isolated atoms in order to enhance the overlapping between them and reinforce the bond.[?]

In molecular calculations using contracted GTOs , the orbital exponents and contraction coefficients of the basis functions are kept fixed at the predetermined values for the basis sets used. Therefore, there is no way for the basis functions to adjust their sizes to differing molecular environments.

One way of improving basis set flexibility is to double the number of orbitals, which is normally known as *Double Z* basis set (DZ). For H this base set contains two s orbitals (1s and 1s), one of which is more contracted than the other. For Li to Ne atoms, there will be four s functions (1s, 1s, 2s, 2s) and two sets of p functions (2p and 2p). The extra flexibility provided by DZ basis set enables the electronic cloud to expand and contract, and take into account dynamic correlation and in particular radial correlation or in-out (whether an electron approaches to the nucleus the others would move away

from the nucleus) and also makes it possible for the effects of anisotropy to be correctly described. Indeed, the electronic cloud expansion or contraction is obtained through the greater or lesser participation of the most diffuse function of each pair. We can improve even more the flexibility by adding another extra set of basis functions: *Triple-Zeta*, and another one (QZ) and so on.

It is clear that this improvement of the basis set involves doubling the number of functions and therefore significantly increasing the computational effort. The aforementioned flexibility is only required in the valence orbitals of the atom, since the characteristics of the internal atoms remain almost unchanged when going from the isolated atom to the atom in the molecule. This led to Pople's group to propose an alternative that was equally flexible but more economical than the double Z basis set. Namely the *split-valence basis method*, which may be described as a valence double zeta method.

The general representation of this approach has the form "x-yzG" where x, y and z are integers defined as follows: x is the number of Gaussian functions whose sum is used to represent each inner-shell basis AO, and yz implies that each valence AO ϕ is to be represented by a sum of *two* AOs (ϕ' and ϕ'', with ϕ' represented by a sum of y Gaussians and ϕ'' by a sum of z Gaussians. The number of "splits" is indicated by the presence of two integers, y and z. If more splits than two are desired, one must indicate so by using more integers; for example, 31 implies two splits, and 311 implies three splits.

1.2.3 Polarization functions

There are some aspects of the bonding between atoms that can only be described if functions are added with larger angular momentum than the ones in the atom ground state. For example, a hydrogen atom in a molecular environment that polarizes it in a preferential direction cannot be described with any of the basis set mentioned above. The reason is that all the functions that are centred on the hydrogen are s orbitals that have spherical symmetry. This polarization could be described if p orbitals were added to hydrogen basis set because the mixing with the s orbitals breaks their spherical symmetry (see figure 42).

These type of functions that have high angular momentum are usually called *polarization functions*. For the atoms of the first row the polarization functions are d orbitals. For transition metals, the polarization functions

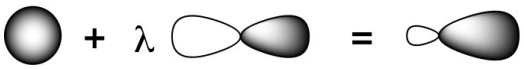


Figure 42: Schematic representation of the mixing between s and p orbitals

are f orbitals and, in general, orbitals that have a higher angular momentum than the occupied ones in the atom.

The polarization functions are generally added to the chosen sp basis set. So, DZP or TZP basis sets would be DZ or TZ basis to which polarization functions have been added. The polarization functions (d) can be added in the heavy atoms and not on hydrogen atoms (6-31G(d) or 6-31G*) or in both kind of atoms (6-31G(d,p) or 6-31G**).

1.2.4 Diffuse functions

For some systems to be properly described, very small exponents (very diffuse functions) must be included in the basis set, in such a way that the maximum of the function, and therefore the maximum of the density probability associated to it, is attained at relatively high values of the electron-nucleus distance. This is the case of the anions, which have an electron that is very weakly linked by the Coulomb field of the nuclei and which is therefore at a much larger distance from them than the other electrons in the system. The situation is similar in weakly bonded systems such as hydrogen-bonded complexes where the distance between the proton donor and the proton acceptor is much larger than in conventional chemical bonds and, for a proper description, the electron density in regions separated from the nuclei has to be accounted for. In general diffuse functions are included as an s function and a set of p functions for each atom of the system. When Pople basis sets are used, the inclusion of diffuse functions is indicated with a symbol +.Or, a 6-31+G(d) basis set is a 6-31G(d) set in which diffuse functions have been added to the heavy atoms. In a 6-31++G(d) basis set diffuse functions have also been added to the hydrogen atoms.

1.3 Effective core potentials

In the study of systems including heavy elements, the size of the atomic basis set increases quickly. However, the orbitals of the most internal shells of these atoms are hardly modified by their participation in the molecular environment because their energies are much lower than those of the orbitals

participating in the bonds. This observation has prompted several authors to define an *Effective Core Potential* (ECP) to represent the effect of the core electrons on the valence ones, preserving the symmetry properties of the atom. Given that the relativistic effects in heavy elements can be considerable, several effective potentials have been developed to include these effects (*relativistic* ECP).

Core potentials can be obtained by least squares fit to accurate *ab initio* calculations. The atomic basis sets to describe the explicit electrons are also given. The size of the core is an important parameter that depends on the type of application, since the polarization of the electrons of the shells close to the valence region can be important. *Large core* ECP include all the electrons except the valence ones and *small core* ECP include in general one shell less.

2 Density Functional Theory

The Density Functional Theory (DFT) represents an alternative to the conventional *ab initio* methods of introducing the effects of electron correlation into the solution to the electronic Schrödinger equation. In 1964, Hohenberg and Kohn showed that the ground state energy depends only on the electron density and thus, they established the basis of the Density Functional Theory (DFT). The only problem is that the precise mathematical formula relating energy to electron density is not known, so it is necessary to resort to approximate expressions.[?]

2.1 Density functions

Let us consider a system of N electrons described by a wavefunction $\Psi(x_1, x_2, , x_N)$. The product $\Psi(x_1, x_2, , x_N)\Psi^*(x_1, x_2, \ldots, x_N)dx_1dx_2 \ldots dx_N$ gives us the probability of finding electron 1 between x_1 and $x_1 + dx_1$, electron 2 between x_2 and $x_2 + dx_2$, , and electron N between x_N and $x_N + dx_N$. The probability of finding electron 1 between x_1 and $x_1 + dx_1$, independently of where the others are found is given by:

$$d\vec{x}_1 \int \Psi(\vec{x}_1, \vec{x}_2, , \vec{x}_N)\Psi^*(\vec{x}_1, \vec{x}_2, , \vec{x}_N)d\vec{x}_2 d\vec{x}_N \tag{37}$$

and taking into account that the electrons are indistinguishable:

$$\rho(\vec{x}_1)d\vec{x}_1 = Nd\vec{x}_1 \int \Psi(\vec{x}_1, \vec{x}_2, , \vec{x}_N)\Psi^*(\vec{x}_1, \vec{x}_2, , \vec{x}_N)d\vec{x}_2 d\vec{x}_N \tag{38}$$

give us the probability of finding an electron between x_1 and $x_1 + dx_1$, independently of where the others are found. $\rho(x)$ is so called *density function*. The *electron density* $\rho(r)$, which can also be obtained experimentally through the X-ray technique, is obtained integrating with respect to the spin coordinate.

$$\rho(\vec{r}_1) = \int \rho(\vec{x}_1)ds_1 = N \int \Psi(\vec{x}_1, \vec{x}_2, , \vec{x}_N)\Psi^*(\vec{x}_1, \vec{x}_2, , \vec{x}_N)ds_1 d\vec{x}_2 d\vec{x}_N \tag{39}$$

When Ψ is normalized:

$$\int \rho(\vec{r})d\vec{r} = N \tag{40}$$

On the other hand, the integral:

$$d\vec{x}_1 d\vec{x}_2 \int \Psi(\vec{x}_1, \vec{x}_2, , \vec{x}_N)\Psi^*(\vec{x}_1, \vec{x}_2, , \vec{x}_N)d\vec{x}_3 d\vec{x}_N \tag{41}$$

give us the probability of finding electron 1 between x_1 and $x_1 + dx_1$ and electron 2 between x_2 and $x_2 + dx_2$, independently of where the others are found. The *second order density*, $\gamma_2(x_1, x_2)$, is defined as:

$$\gamma_2(\vec{x}_1, \vec{x}_2) = N(N-1) \int \Psi(\vec{x}_1, \vec{x}_2, , \vec{x}_N)\Psi^*(\vec{x}_1, \vec{x}_2, , \vec{x}_N)d\vec{x}_3 d\vec{x}_N \tag{42}$$

where N(N-1) are all the possible electrons pairs that can be formed, and given that the electrons are indistinguishable, $\gamma_2(x_1, x_2)dx_1 dx_2$ consequently gives us the probability of finding any electron between x_1 and $x_1 + dx_1$ and another between x_2 and $x_2 + dx_2$. Integrating with respect to the spin variables, we obtain the *two-electron density* or *pair function*:

$$\gamma_2(\vec{r}_1, \vec{r}_2) = \int \gamma_2(\vec{x}_1, \vec{x}_2)ds_1 ds_2 \tag{43}$$

which gives us the probability of finding any two electrons, one between r_1 and r_1+dr_1 and the other between r_2 and r_2+dr_2, under any spin combination $(\alpha\alpha, \alpha\beta, \beta\alpha, \beta\beta)$.

2.2 Exchange-correlation hole

Two-electron density $\gamma_2(r_1, r_2)$, gives us the probability density of simultaneously finding two electrons, one in r_1 and the other in r_2. In other words, the two-electron density contains information about the correlated motion of two electrons. The same information about the electron correlation is contained in the exchange-correlation density, the conditional probability density, and the exchange-correlation holes described below.

The exchange correlation density, $\Gamma_{XC}(r_1, r_2)$, is defined as follows:

$$\gamma_2(\vec{r}_1, \vec{r}_2) = \rho(\vec{r}_1)\rho(\vec{r}_2) + \Gamma_{XC}(\vec{r}_1, \vec{r}_2) \tag{44}$$

Considering electrons as uncorrelated independent particles, the term $\rho(r_1)\rho(r_2)$ is the probability density of finding an electron in r_1 and another in r_2. Consequently, the exchange-correlation density, $\Gamma_{XC}(r_1, r_2)$, represents the difference between the probability density of finding two electrons, one in r_1 and the other in r_2, correlated or uncorrelated.

If we divide equation (44) by $\rho(r_1)$ we obtain the conditional probability density, $\mathrm{P}(\mathrm{r}_1,\mathrm{r}_2)$:

$$\frac{\gamma_2(\vec{r}_1, \vec{r}_2)}{\rho(\vec{r}_1)} = P(\vec{r}_1, \vec{r}_2) \tag{45}$$

$\mathrm{P}(\mathrm{r}_1,\mathrm{r}_2)$ gives us the probability of finding an electron in r_2 when there is already another one in r_1. It can be more clearly seen if we rearrange 45 in the following way:

$$\gamma_2(\vec{r}_1, \vec{r}_2) = \rho(\vec{r}_1)P(\vec{r}_1, \vec{r}_2) \tag{46}$$

If we divide equation (44) by $\rho(r_1)$ we obtain:

$$P(\vec{r}_1, \vec{r}_2) = \rho(\vec{r}_2) + \rho_{XC}(\vec{r}_1, \vec{r}_2) \tag{47}$$

where $\rho_{XC}(\vec{r}_1, \vec{r}_2)$ is given by:

$$\rho_{XC}(\vec{r}_1, \vec{r}_2) = \frac{\Gamma_{XC}(\vec{r}_1, \vec{r}_2)}{\rho(\vec{r}_1)} \tag{48}$$

$\rho_{XC}(r_1, r_2)$ is called the *exchange-correlation* or *Fermi-Coulomb hole*. According to equation (47), the exchange correlation hole is the correction term, which must be added to the unconditional probability in order to obtain the

conditional probability. It is the region around the electron in which the presence of other electrons is excluded to a greater or lesser degree. Substituting (47) into (46) we have:

$$\gamma_2(\vec{r}_1,\vec{r}_2) = \rho(\vec{r}_1)\rho(\vec{r}_2) + \rho(\vec{r}_1)\rho_{XC}(\vec{r}_1,\vec{r}_2) \tag{49}$$

Because the total electron density is the sum of densities α and β, it is possible to separate the different contributions into $\rho_{XC}(r_1,r_2)$. Thus:

$$\gamma_2^{\alpha\alpha}(\vec{r}_1,\vec{r}_2) = \rho^{\alpha}(\vec{r}_1)\rho^{\alpha}(\vec{r}_2) + \Gamma_{XC}^{\alpha\alpha}(\vec{r}_1,\vec{r}_2) \tag{50}$$

$$\begin{aligned}\frac{\gamma_2^{\alpha\alpha}(\vec{r}_1,\vec{r}_2)}{\rho^{\alpha}(\vec{r}_1)} = P^{\alpha\alpha}(\vec{r}_1,\vec{r}_2) &= \rho^{\alpha}(\vec{r}_2) + \frac{\Gamma_{XC}^{\alpha\alpha}(\vec{r}_1,\vec{r}_2)}{\rho^{\alpha}(\vec{r}_1)} \\ &= \rho^{\alpha}(\vec{r}_2) + \gamma_{XC}^{\alpha\alpha}(\vec{r}_1,\vec{r}_2)\end{aligned} \tag{51}$$

where $\gamma_{XC}^{\alpha\alpha}(r_1,r_2)$ is the *Fermi hole* and represents for a specific reference electron with α spin located in r_1, the region of the space from which the presence of another electron of the same spin is excluded to a greater or lesser extent. The probability of finding an electron in r_2 when there is another of the same spin in r_1 is reduced, especially for small $|r_1 - r_2|$. In addition:

$$\frac{\gamma_2^{\alpha\beta}(\vec{r}_1,\vec{r}_2)}{\rho^{\alpha}(\vec{r}_1)} = P^{\alpha\beta}(\vec{r}_1,\vec{r}_2) = \rho^{\beta}(\vec{r}_2) + \frac{\Gamma_{XC}^{\alpha\beta}(\vec{r}_1,\vec{r}_2)}{\rho^{\alpha}(\vec{r}_1)} = \rho^{\beta}(\vec{r}_2) + \gamma_{XC}^{\alpha\beta}(\vec{r}_1,\vec{r}_2) \tag{52}$$

$$\tag{53}$$

where $\rho_{XC}^{\alpha\beta}(r_1,r_2)$ is called the *Coulomb hole* and would give us the zone where the presence of a β electron is excluded or favoured in the case in which we have a reference electron with α spin located at r_1.

It can be demonstrated that

$$\int \rho_{XC}(\vec{r}_1,\vec{r}_2)d\vec{r}_2 = -1 \tag{54}$$

which constitutes the so-called *sum rule*. Thus an electron located in r creates around itself a hole, a charge deficit, with the displaced charge being exactly equal to that of a positive electron. This hole follows the motion of the electron.

2.3 Expected value expression

The expected value of the electron energy can be expressed for a system of n electrons and N nuclei for a given nuclear configuration as a function of

the first-order density matrix and of the two-electron density. The operator which must be taken into consideration is the electronic Hamiltonian, $\hat{H}_{el}$, by which:

$$E_{el} = \langle \Psi \left| \hat{H}_{el} \right| \Psi \rangle = \int_{r_1=r_1} \left(-\frac{1}{2}\nabla^2(1) + \hat{v}(1) \right) \rho_1(\vec{r}_1, \vec{r}_1) d\vec{r}_1 + \\ + \frac{1}{2} \int \int \frac{\gamma_2(\vec{r}_1, \vec{r}_2)}{|\vec{r}_1 - \vec{r}_2|} d\vec{r}_1 d\vec{r}_2 \qquad (55)$$

substituting equation (49) in the previous equation we are left with:

$$E_{el} = \int_{r_1=r_1} \left(-\frac{1}{2}\nabla^2(1) \right) \rho_1(\vec{r}_1, \vec{r}_1) d\vec{r}_1 + \int \hat{v}(1)\rho_1(\vec{r}_1) d\vec{r}_1) + \\ + \frac{1}{2} \iint \frac{\rho(\vec{r}_1)\rho(\vec{r}_2)}{|\vec{r}_1 - \vec{r}_2|} d\vec{r}_1 d\vec{r}_2 + \frac{1}{2} \iint \frac{\rho(\vec{r}_1)\rho_{XC}(\vec{r}_1, \vec{r}_2)}{|\vec{r}_1 - \vec{r}_2|} d\vec{r}_1 d\vec{r}_2 \qquad (56)$$

The different terms appearing in equation (56) are, respectively, the kinetic energy of the electrons, the electron-nuclear potential energy, the Coulomb electron repulsion and the electronic exchange-correlation energy. For this last term it is possible to write:

$$W_{XC}[\rho] = \frac{1}{2} \int \rho(\vec{r}_1) d\vec{r}_1 \int \frac{\rho_{XC}(\vec{r}_1, \vec{r}_2)}{|\vec{r}_1 - \vec{r}_2|} d\vec{r}_2 = \int \rho(\vec{r}_1) \mu_{XC}^{hole}(\vec{r}_1) d\vec{r}_1 \qquad (57)$$

W_{XC} represents the interaction of $\rho(r_1)$ with its exchange-correlation hole $\rho_{XC}(r_1, r_2)$. The electron interacts with an effective charge distribution corresponding to a positive electron, as follows from the sum rule.

Equation (56) shows that it is possible to express the energy from the first- and second-order density functions and matrices and the energy is consequently said to be a functional of the density. Understanding by functional a mathematical expression associating a number to a function., i.e., a functional is a function whose argument is also a function.

2.4 The Hohenberg-Kohn theorems

2.4.1 The first theorem

Equation (56) shows that it is possible to express the electron energy of a system as a functional of first- and second-order density functions and matrices. The first Hohenberg-Kohn theorem,[?] which is considered to have given

rise to the rigorous DFT, goes further to demonstrate that:

Any observable of a stationary non-degenerate ground state can be calculated, exactly in theory, from the electron density of the ground state. In other words, any observable can be written as a functional of the electron density of the ground state.

For the cases where this theorem is valid we have that:

$$E[\rho] = T[\rho] + V_{Ne}[\rho] + V_{ee}[\rho](+V_{NN}) \tag{58}$$

T[ρ] and $V_{ee}[\rho]$ are universal functional, given that they do not depend on the external potential. They are usually encompassed within the Hohenberg-Kohn functional $F_{HK}[\rho]$, with which:

$$E_\mu[\rho] = \int \rho(\vec{r})\mu(\vec{r})d\vec{r} + F_{HK}[\rho] \tag{59}$$

where $E_\mu[\rho]$ indicates that, for a specific external potential v(r), the energy is a functional of density.

2.4.2 The second theorem

The second Hohenberg-Kohn theorem provides the variational principle for E[ρ]. It can be expressed as follows:

The electron density of a non-degenerate ground state can be calculated, exactly in theory, determining the density that minimizes the energy of the ground state.

Or what is the same, for a trial density, $\widetilde{\rho}(\vec{r})$, that is v-representable and N-representable it is fulfilled that:

$$E_0 \leq E_\mu\left[\widetilde{\rho}(\vec{r})\right] \tag{60}$$

With this equation we obtain a variational principle for the energy in the framework of the DFT. This principle assures that any trial density results in an energy greater or equal to the exact energy of the ground state. Therefore , to obtain the exact density of the ground state, we will have to find the density that minimizes the energy:

$$[\frac{\delta E_\mu[\rho]}{\delta\rho}] = 0 \tag{61}$$

2.5 Fundamental equation

The minimisation of the energy functional with respect to the electron density, $\delta E_\mu[\rho] = 0$, must be carried out assuring the conservation of the N-representability during the optimization process. This is done by introducing the restriction $\int \rho(\vec{r})d\vec{r} - N = 0$ by means of Lagrange undetermined multipliers method. For that the following function is constructed:

$$E_\mu[\rho] - \nu \left[\int \rho(\vec{r})d\vec{r} - N \right] \tag{62}$$

where ν is the undetermined Lagrange multiplier and has the meaning of a chemical potential. Next it is minimized:

$$\delta \left[E_\mu[\rho] - \nu \left(\int \rho(\vec{r})d\vec{r} - N \right) \right] = 0 \tag{63}$$

and proceeding from the definition of a functional differential we are left with:

$$\int \frac{\delta E_\mu[\rho]}{\delta \rho(\vec{r})} \delta\rho(\vec{r})d\vec{r} - \nu \int \delta\rho(\vec{r})d\vec{r} = 0 \tag{64}$$

and rearranging:

$$\int \left\{ \frac{\delta E_\mu[\rho]}{\delta \rho(\vec{r})} - \nu \right\} \delta\rho(\vec{r})d\vec{r} = 0 \tag{65}$$

which provides the condition of constrained minimisation and allows the minimum value of ν to be obtained:

$$\nu = \frac{\delta E_\mu[\rho]}{\delta \rho(\vec{r})} = \mu_n(\vec{r}) + \frac{\delta F_{HK}[\rho]}{\delta \rho(\vec{r})} \tag{66}$$

Equation (66) is known as *fundamental equation of density functional theory.*

2.6 The Kohn and Sham method

Equation (66) gives us a formula to minimise the energy and thereby determine the density of the ground state. The problem that arises is that the exact expression relating F_{HK} with the density is unknown. In particular, the exact form of $\mathrm{T}[\rho]$ is not known precisely enough. On the other hand, the kinetic energy is easily calculated if Ψ is known. Proceeding from the

expression of T[Ψ] for a single determinant function, Kohn and Sham in 1965 proposed an ingenious method to calculate the energy from ρ.[?]

In 1975 Gilbert[?] proved that for any electron density normalized to N electrons there is a finite set of one-electron functions $\{\varphi_i\}$ that:

$$\rho(r) = \sum_i \varphi_i^*(r)\varphi_i(r) \tag{67}$$

known as the Gilbert decomposition of the electronic density. Kohn and Sham considered these one-electron functions as orbitals, establishing analogies between DFT and the wavefunction based methods. They used as a *reference system* a system of N electrons that do not interact among themselves and that move under a effective potential created by the other electrons, as in the Hartree-Fock model. Thus, they introduced this finite set of orbitals $\{\varphi_i\}$ where each orbital describes one of the no interacting electrons. Within this assumption, an analogue to the HF functional for the kinetic energy can be proposed:

$$T_s[\rho] = \sum_i \langle \varphi \left| -\frac{1}{2}\nabla^2 \right| \varphi \rangle \tag{68}$$

where $T_s[\rho]$ does not represent the exact kinetic energy but an approximation, due to the fact that the functions φ_i are not the orbitals that make up the wavefunction. To avoid confusions, from now on we will call the set $\{\varphi_i\}$ Kohn-Sham (KS) orbitals.

Now we can express the total energy of the system as:

$$E[\rho] = T_s[\rho] + \frac{1}{2}\int \frac{\rho(r_1)\rho(r_2)}{r_{12}} dr_1 dr_2 + \int \rho(r) V_{eN} dr + E_{XC}[\rho] \tag{69}$$

Thus the effective potential within the electrons are moving in the Kohn-Sham model is given by:

$$\mu_{eff}(r) = \frac{1}{2}\int \frac{\rho(r_2)}{r_{12}} dr_2 + V_{eN} + \frac{\delta E_{XC}[\rho]}{\delta \rho(r)} \tag{70}$$

where the last terms represents the exchange-correlation potential

$$\mu_{XC} = \frac{\delta E_{XC}[\rho]}{\delta \rho(r)} \tag{71}$$

Now we are able to construct the Hamiltonian:

$$\hat{H}_s = \sum_i \left[-\frac{1}{2}\nabla_i^2 + \mu_{eff} \right] \tag{72}$$

As in the Hartree-Fock approximation, we can construct a Slater determinant with the set of M KS orbitals $\{\varphi_i\}$. Therefore, the applications of the Hamiltonian $\hat{H}_s$ to a such a wavefunction and minimize the energy is the same as solve the M uncoupled equations:

$$\left[-\frac{1}{2}\nabla_i^2 + \mu_{eff}(r) \right] \varphi_i^{KS} = \epsilon_i \varphi_i^{KS} \tag{73}$$

that is known as the *Kohn-Sham equations*. As it happened in the Hartree-Fock model, the effective potential depends on the electronic density eq. (70) , so it is needed an initial set of KS orbitals $\{\varphi_i^{(0)}\}$ that would be improved iteratively. The only unknownquantity from these equations is the exchange-correlation potential, if the exact form was known, we would get the exact value for the energy of the ground state of the system via the Kohn-Sham equations, because this treatment does not include any approximations. The only approximation made in the DFT method is the expression of μ_{XC}. Thus it is expected that the electron density $\rho(r)$ will approach the exact density as $\mu_{XC}(r)$ approaches the exact exchange-correlation energy. In this way, DFT has the potential capacity to incorporate all the correlation energy, unlike HF. These are the summarized steps of the algorithm to solve equation (73):

Step 1: Definition of a first set of trial orbitals $\{\varphi_i^{(0)}\}$.

Step 2: Propose a functional for the exchange-correlation energy and calculate the exchange-correlation potential as in equation (71).

Step 3: Construct the effective potential as given by equation (70).

Step 4: Solve the Kohn-Sham equations (73) in order to obtain an improved set of KS orbitals: $\{\varphi_i^{(1)}\}$.

Step 5: With the new set of improved KS orbitals, a new density $\rho^{(1)}$ is constructed eq. (67). If the calculation has converged (with some pre-established criteria as in the HF method), the approximate value for the ground state energy is obtained by means of (69). If the calculation has not converged, we return to the 3^{rd} step.

2.7 Approximations to the exchange-correlation potential

A good exchange-correlation potential is a basic aspect of the DFT. Good approximations to the exchange-correlation energy are required to put this theory into practice. The simplest but at the same time tremendously useful approximation is the local density approximation(LDA). At the following level the so-called non-local or generalized gradient approximations (GGA) are found. These two types of approximations together with the meta-GGZ and hybridfunctionals are discussed below.

2.7.1 Local density approximation (LDA)

Hohenberg and Kohn showed that if ρ varies extremely slowly with position, then $\mathrm{E}_{XC}[\rho]$ is accurately given by

$$E_{XC}^{LDA}[\rho] = \int \rho(r)\epsilon_{XC}(\rho)dr \tag{74}$$

where the integral is over all space, dr stands for dx dy dz and $\epsilon_{XC}(\rho)$ is the exchange plus correlation energy per electron in a homogeneous electron gas with electron density ρ. *Jellium* as a hypothetical electrically neutral, infinite-volume system consisting of an infinite number of interacting electrons moving in a space throughout which positive charge is continuously and uniformly distributed. The number of electrons per unit volume in the jellium has a nonzero constant value ρ. The electrons in the jellium constitute a homogeneous electron gas. Taking the functional derivative of E_{XC}^{LDA}, we find:

$$\mu_{XC}^{LDA} = \frac{\delta E_{XC}^{LDA}}{\delta\rho} = \epsilon_{XC}(\rho(r)) + \rho(r)\frac{\delta\epsilon_{XC}(\rho)}{\delta\rho} \tag{75}$$

Kohn and Sham suggested the use of equations (74) and (75) as approximations to E_{XC} and μ_{XC}. This procedure is called the *local density approximation* (LDA). One can show that ϵ_{XC} can be written as the sum of exchange and correlation parts:

$$\epsilon_{XC}(\rho) = \epsilon_X(\rho) + \epsilon_c(\rho) \tag{76}$$

where expression for the correlation energy as a function of the electronic density was found by Slater[?] to be

$$\epsilon_x(\rho) = -\frac{3}{4}\left(\frac{3}{\pi}\right)^{\frac{1}{3}}(\rho(r))^{\frac{1}{3}} \tag{77}$$

Thus the exchange functional is known as the Slater exchange and is abbreviated with an S. The correlation part ϵ_C has been calculated and the results have been expressed as a very complicated function ϵ_C^{VWN} of ρ by Vosko, Wilk and Nusair[?](VWN). Thus

$$\epsilon_C(\rho) = \epsilon_C^{VWN}(\rho) \tag{78}$$

where ϵ_C^{VWN} is known function. Taking this into account, a E_{XC}^{LDA} would be the SVWN, that indicates that the exchange part is given by the Slater functional while the correlation part is describe by the Vosko, Wink and Nusair functional.

2.7.2 The Local-Spin-Density Approach (LSDA)

For open-shell molecules and molecular geometries near dissociation, the *local-spin-density approximation* gives better results than the LDA. Whereas in the LDA, electrons with opposite spins paired with each other have the same spatial KS orbitals, the LSDA allows such electrons to have different spatial KS orbitals $\varphi_{i\alpha}^{KS}$ and $\varphi_{i\beta}^{KS}$. The LSDA is thus the analogous to the UHF method.

As in the UHF method, allowing differing KS orbitals for electrons with different spins can produce a wavefunction for the reference system s that is not an eigenfunction of $\hat{S}^2$, but this spin contamination is less of a problem in KS DFT than in the UHF method.

Despite the fact that ρ in a molecule is not a slowly varying function of position, the LSDA works surprisingly well for calculating molecular equilibrium geometries, vibrational frequencies, and dipole moments, ever for transition-metal compounds, where Hartree-Fock calculations often give poor results. However, calculated LSDA molecular atomization energies are very inaccurate. Accurate dissociation energies require functional that go beyond LSDA.[?]

2.7.3 Gradient-Corrected (GGA) Functionals

The LDA and LSDA are based on the uniform-electron-gas model, which is appropriate for a system where rho varies slowly with position. The integrand in the expression (74) for E_{XC}^{LDA} is a function of only ρ, and the integrand in E_{XC}^{LSDA} is a function of only ρ^α and ρ^β. Functionals that go beyond LSDA

aim to correct the LSDA for the variation of electron density with position. A common way to do this is by including the gradients of ρ^α and ρ^β in the integrand. Thus

$$E_{XC}^{GGA}[\rho^\alpha, \rho^\beta] = \int f(\rho^\alpha(r), \rho^\beta(r), \nabla\rho^\alpha(r), \nabla\rho^\beta(r))dr \tag{79}$$

where f is some function of the spin densities and their gradients. The letters GGA stand for *generalized-gradient approximation*. The term *gradient-corrected functional* is also used. E_{XC}^{GGA} is usually split into exchange and correlation parts, which are modelled separately:

$$E_{XC}^{GGA} = E_X^{GGA} + E_C^{GGA} \tag{80}$$

Approximate gradient-corrected exchange and correlation energy functionals are developed using theoretical considerations such as the known behaviour of the true (but unknown) functionals E_X and E_C in various limiting situations as a guide, with often some empiricism thrown in, by choosing the values of parameters in the functionals to give good performance for known values of various molecular properties.

Some commonly used GGA exchange functionals E_X are Perdew and Wangs 1986 functional (which contains non empirical parameters), designated by PW86 or PWx86, Beckes 1988 functional,[?] denoted B88,Bx88, Becke88, or B, and Perdew and Wangs 1991 exchange functional[?]PWx91. The explicit form of the B88 exchange functional is

$$E_X^{B88} = E_X^{LSDA} - b\sum_{\sigma=\alpha,\beta}\int \frac{(\rho^\sigma)^{\frac{4}{3}}\chi_\sigma^2}{1 + 6b\chi_\sigma \ln\left[\chi_\sigma + (\chi_\sigma^2 + 1)^{\frac{1}{2}}\right]}dr \tag{81}$$

where $\chi_{sigma} \equiv |\nabla\rho^\sigma|/(\rho^{\sigma\,\frac{4}{3}})$, b is an empirical parameter whose value 0.0042 atomic units was determined by fitting known Hartree-Fock exchange energies of several atoms, and

$$E_X^{LSDA} = -\frac{3}{4}\left(\frac{6}{\pi}\right)^{\frac{1}{3}}\int[(\rho^\alpha)^{\frac{4}{3}} + (\rho^\beta)^{\frac{4}{3}}]dr \tag{82}$$

The PWx86 functional (which has non empirical parameters) and the B88 exchange functional work about equally well in predicting molecular properties.[?]

Commonly used GGA correlation functionals E_C include the Lee-Yang-Parr (LYP) functional,[?] the Perdew 1986 correlation functional (P86 or

Pc86), and the Perdew-Wang 1991 parameter-free correlation functional (PW91 or PWc91).

The Perdew-Burke-Ernzerhof (PBE) exchange and correlation functional[?] has no empirical parameters.

Any exchange functional can be combined with any correlation functional. For example, the notation BLYP/6-31G* denotes a density functional calculation done with the Becke 1988 exchange functional and the Lee-Yang-Parr correlation functional, with the KS orbitals expanded in a 6-31G* basis set.

2.7.4 Meta-GGA Functionals

The GGA density functionals of the form 79 depend on the ground state electron probability density ρ and its first derivatives. One way to improve on GGA functionals is to go to functionals that also depend on the second derivatives of ρ and/or a quantity called the kinetic-energy density. Such functionals are called *meta-GGA* functionals and have the form:

$$E_{XC}^{MGGA}[\rho^\alpha, \rho^\beta] = \int f(\rho^\alpha, \rho^\beta, \nabla\rho^\alpha, \nabla\rho^\beta \nabla^2\rho^\alpha, \nabla^2\rho^\beta \tau_\alpha \tau_\beta) dr \tag{83}$$

where the *Kohn-Sham kinetic-energy density* for the spin α electrons is defined by

$$\tau_\alpha \equiv \frac{1}{2}\sum_i |\nabla\theta_{i\alpha}^{KS}|^2 \tag{84}$$

2.7.5 Hybrid Functionals

Hybrid exchange-correlation functionals are widely used. A hybrid functional mixes together the formula (85) for E_X with GGA (or meta-GGA) E_X and E_C formulas.

$$E_X \equiv -\frac{1}{4}\sum_{i=1}^{n}\sum_{j=1}^{n}\langle\theta_i^{KS}(1)\theta_j^{KS}(2)|1/r_{12}|\theta_j^{KS}(1)\theta_i^{KS}(2)\rangle \tag{85}$$

In (85), E_X is defined by the same formula used for the exchange energy in Hartree-Fock theory, except that the Hartree-Fock orbitals are replaced by the Kohn-Sham orbitals. Since, in practice, KS orbitals are found to rather closely resemble the HF orbitals, the DFT exchange energy so computed is

close to the Hartree-Fock exchange energy.

One example of hybrid functional is the popular B3LYP (or Becke3LYP) hybrid GGA functional (where the 3 indicates a three-parameter functional). It is defined by

$$E_{XC}^{B3LYP} = (1 - a_0 - a_x)E_X^{LSDA} + a_0 E_X^{exact} + a_x E_X^{B88} + (1 - a_c)E_C^{VWN} + a_C E_C^{LYP} \quad (86)$$

where E_X^{exact} (which often is denoted E_X^{HF}), since it uses Hartree-Fock definition of E_X) is given by 85, and where the parameter values $a_0 = 0.20, a_x = 0.72$ and $a_c = 0.81$ were chosen to give good fits to experimental molecular atomization energies.

The GGA, meta-GGA, hybrid-GGA and hybrid-meta-GGA functionals give not only good equilibrium geometries, vibrational frequencies and dipole moments, but also generally accurate molecular atomization energies.

2.7.6 Evaluation of Functionals

There is not correct answer to the question Which is the best DFT functional? Since one finds that functionals that give good results for organic compounds may give inferior results for inorganic compounds, and functionals that give good results for energy changes in reactions may give inferior results for activation energies of reactions. Thus the best functional to use depends on the kinds of compounds being studied and on which properties are being calculated.

2.8 Advantages and disadvantages of DFT

The main advantage of DFT is that allows for correlations effects to be included in a calculation that takes roughly the same time as a Hartree-Fock calculation, which does not include correlation. However, despite it success DFT is not a panacea and it has some drawbacks and failings, some of them are that:

- Due to the fact that the Hohenberg-Kohn-Sham theory is basically a ground-state theory it fails in general when dealing with electronic excited states.
- Because approximate functionals are used, KS DFT is not variational and can yield an energy below the true ground-state energy.

- The true E_{XC} contains a self-interaction correction that exactly cancels the self-interaction energy in $\frac{1}{1}\iint \rho(r_1)\rho(r_2)r_{12}^{-1}dr_1dr_2$, but most currently used functionals are not completely free of self-interaction.

- Many of the currently used E_XC functionals fail for van der Waals molecules and systems weakly bonded.

3 G4 theory

A number of approaches, based on quantum chemical methods, have been developed over the past decade and a half to make accurate predictions of thermochemical data. The Gaussian-n (Gn) theories (n = 1, 2, 3),[92–95] employ a set of calculations with different levels of accuracy and basis sets with the goal of approaching the exact energy. In the Gn approach, a high level correlation calculation [e.g., QCISD(T) adn CCSD(T)] with a moderate sized basis set is combined with energies from lower level calculations (e.q. MP4 and MP2) with larger basis sets to approximate the energies of more expensive calculations. In addition, several molecule-independent empirical parameters [higher level correction (HLC) terms] are included to estimate remaining deficiencies, assuming that they are systematic. Therefore, the Gaussian-4 theory[96] (G4 theory) is a composite technique aimed at getting accurate energies without requiring extensive computer resources. This approach depends on a cancellation of errors as well as well-defined parameters to achieve this.

In the G4 theory, a sequence of well-defined *ab initio* molecular orbital calculations is performed to arrive at a total energy for a given molecular species. This method allows the calculation of energies of molecular species containing first-row (Li-F), second row (Na-Cl), and third-row main group elements. The steps in the G4 theory are as follows.

1. The equilibrium structure is obtained at the B3LYP/6-31G(2df,p) level. Spin-restricted theory is used for singlet states and spin-unrestricted theory for others.

2. The B3LYP/6-31G(2df,p) equilibrium structure is used to calculate harmonic frequencies, which are then scaled by a factor of 0.9854[97] to take into account of known deficiencies at this level. These frequencies give the zero-point energy $E(ZPE)$ used to obtain E_0.

3. The Hartree-Fock energy limit $E(HF/limit)$ is calculated. The Hartree-Fock basis set limit is determined using a linear two-point extrapolation scheme[98,99] and Dunning's aug-cc-pVnZ basis sets,[100–102]

$$E_{\text{HF/aug-cc-pV}n\text{Z}} = E_{textHF/limit} + B\exp(-\alpha n). \tag{87}$$

where n is the number of contractions in the valence shell of the basis set and α is an adjustable parameter. The authors found that calculating the HF limit using n=4, and $n + 1 = 5$ (aug-cc-pVQZ and aug-cc-pV5Z basis sets), and $\alpha = 1.63$ gave nearly converged values for a set of large molecules from the G3/05 test set. In order to reduce the computational time required, the authors also modified the standard aug-cc-pVQZ and aug-cc-pV5Z basis sets by reducing the number of diffuse functions on heavy atoms and by reducing the hydrogen basis set as described in the paper.[96] These modifications save significant computer time, without reducing the accuracy.

4. A series of single point correlation energy calculations is then carried out. The first is based on the complete fourth-order Moller-Plesset perturbation theory with the 6-31G(d) basis set, MP4/6-31G(d). This is modified by corrections from additional calculations.

   (a) a correction for diffuse functions,

$$\Delta E(+) = E[\text{MP4/6-31+G}(d)] - E[\text{MP4/6-31G}(d)], \tag{88}$$

   (b) a correction for higher polarization functions,

$$\Delta E(2df,p) = E[\text{MP4/6-31G}(2df,p)] - E[\text{MP4/6-31G}(d)], \tag{89}$$

   (c) a correction for correlation effects beyond a fourth-order perturbation theory using a coupled cluster theory,

$$\Delta E(\text{CC}) = E[\text{CCSD(T)/6-31G}(d)] - E[\text{MP4/6-31G}(d)], \tag{90}$$

   (d) a correction for larger basis set effects and for teh nonadditivity caused by the assumption of separate basis set extensions for diffuse functions and higher polarization functions,

$$\begin{aligned} \Delta E(\text{G3LargeXP}) &= E[\text{MP2(full)/G3LargeXP}] \\ &- E[\text{MP2/6-31G}(2df,p)] \\ &- E[\text{MP2/6-31+G}(d)] \\ &+ E[\text{MP2/6-31G}(d)]. \end{aligned} \tag{91}$$

The MP4 and CCSD(T) calculations are done in the frozen core approximation, while the MP2 calculation with the large basis set is done with all electrons correlated.

5. The MP4/6-31G(d) energy and the four correlation corrections from step 4 are combined in an additive manner along with a correction for the HF limit (step 3) and a spin-orbit correction, ΔE(SO)

$$\begin{aligned} \Delta E(\text{combined}) &= E[\text{MP4/6-31G}(d)] + \Delta E(+) \\ &+ \Delta E(2df,p) + \Delta E(\text{CC}) \\ &+ \Delta E(\text{G3LargeXP}) + \Delta E(\text{HF}) \\ &+ \Delta E(\text{SO}). \end{aligned} \tag{92}$$

The ΔE(HF) is calculated as the difference between E(HF/G3LargeXP) calculated in step 4 and E(HF/limit) calculated in step 3[*i.e.*, ΔE(HF) = E(HF/limit) - E(HF/G3LargeXP)].

6. A high level correction (HLC) is added to take into account to take into account the remaining deficiencies in the energy calculations,

$$E_e(\text{G4}) = E(\text{combined}) + E(\text{HLC}). \tag{93}$$

where the HLC form is

$$E(\text{HLC}) = \begin{cases} -An_\beta & \text{closed shell molecules} \\ -A'n_\beta - B(n_\alpha - n_\beta) & \text{open shell systems} \\ -Cn_\beta - D(n_\alpha - n_\beta) & \text{for atoms (including atomic ions)} \end{cases} \tag{94}$$

The n_β and n_α are the number of β and α valence electrons, respectively, with $n_\alpha \geq n_\beta$. The A, A', B, and C values are chosen to give the smallest average absolute deviation from experiment for the G3/05 test set. The values obtained are, A = 6.947 mhartree, B = 2.441 mhartree, C = 7.116 mhartree, D = 1.414 mhartree, and A' = 7.128 mhartree.

7. Finally, the total energy at 0 K is obtained by adding the zero-point energy, obtained from the frequencies of step 2, to the total energy,

$$E_0(\text{G4}) = E_e(\text{G4}) + E(\text{ZPE}). \tag{95}$$

The energy E_0 is referred to as the "G4-energy". The final total energy obtained with such procedure is effectively at the CCSD(T,full)/G3LargeXP + HFlimit level.

The G3/05 test set[103] was used to assess the performance of the G4-theory.[96] This test set contains thermochemical data chosen based on a listed accuracy of ± 1 kcal mol^{-1}or better. The G3/05,[103] contains 270 enthalpies of formation, 105 ionization energies, 63 electron affinities, 10 proton affinities, and 6 hydrogen-bonded complexes. The overall absolute deviation for the 454 energies in the G3/05 test set is 0.83 kcal mol^{-1}.

4 Population analysis methods

The previous sections were devoted to different methods of finding a wavefunction, more or less precise, by means of different approximations to the Schrödinger equation. By solving the Schrödinger equation the energy of the system is obtained, and the geometry, but there is much more information contained it the wavefunction. In particular, appropriate auxiliary tools allowing electron pair localization[104,105] have been long pursued in quantum chemistr to explain the nature of the chemical bond[106–108] and with the aim of establishing a link between the rigorous but abstract wavefunction and the classical chemical concepts based on the Lewis theory**??** and the valence shell electron pair repulsion (VSEPR) model of molecular geometry.[109,110] In this section we wil briefly described three of these methods used during this work, namely, the natural bond orbital (NBO) method, the Quantum Theory of Atoms in Molecules (QTAIM), and the electron localization function (ELF).

4.1 Populations analysis of Natural Bond Orbitals

On a wide range of quantum-chemist studies it is very important to determine the electronic configuration and the net charge of each atom, namely the electronic distribution within the molecule, due to the great utility of these properties to understand the chemical processes.
On a wide range of quantum-chemist studies it is very important to determine the electronic configuration and the net charge of each atom, namely the electronic distribution within the molecule, due to the great utility of these properties to understand the chemical processes.

Over all the several approaches proposed to analyse the atomic population, the developed by Mulliken[?] has been the most widely used to characterize the charge distribution hitherto. However, it has been showed repeatedly that the Mulliken method fails when it concerns to the calculation of the

atomic charges and orbitals population. Some of the drawbacks are: sometimes it yields to results with no physical meaning, its highly dependence on the basis set and the low reliability of the charge distributions obtained with this method when dealing in compounds with highly ionic character.

These limitations of the Mulliken method have encourage the search of new method to analyse the populations that overcome these disadvantages.

A quite popular method for population analysis is based on Natural Orbitals (NO). These orbitals are the eigenvectors of the first-order reduced density matrix (96), whereas the population (occupancy) numbers are the eigenvalues.

$$\gamma_i(r_1, r_1) = N_{elec} \int \Psi^*(r_1, r_{N_{elec}}) \int \Psi^*(r_1, r_{N_{elec}}) dr_{N_{elec}} \tag{96}$$

Weinhold and co-workers,[?] used this Natural Orbital definition to distributed the electrons between atomic and molecular orbitals and thus obtain the natural charges and the molecular bonds. Let us suppose that the basis functions are arranged in such a way that the functions corresponding to the A atom are found first that the ones of the B atom and so forth:

$$\chi_1^A, \chi_2^A, \chi_3^A, \ldots, \chi_1^B, \chi_2^B, \chi_3^B, \ldots, \chi_1^C, \chi_2^C, \chi_3^C, \ldots \tag{97}$$

so now, the density matrix , D, can be written as basis set functions blocks belonging to an specific centre:

$$D = \begin{pmatrix} D^{AA} & D^{AB} & D^{AC} & \vdots \\ D^{AB} & D^{BB} & D^{BC} & \vdots \\ D^{AC} & D^{BC} & D^{CC} & \vdots \\ \ldots & \ldots & \ldots & \ddots \end{pmatrix} \tag{98}$$

The A *natural atomic orbitals* (NAOs) are obtained by diagonalizing the D^{AA} block, the B NAOs are obtained by diagonalizing the D^{BB} and so on. In general, the NAOs are not orthogonal and therefore, their populations does not sum up to the total number of electrons, hence it is necessary to make them orthogonal.
The NAOs are automatically ordered in importance by occupancy. Consistent with chemical intuition, only the core and valence shell NAOs are found to have significant occupancies, compared to the extra-valence Rydberg-type

NAOs that complete the span of the basis. The effective dimensionality of the NAO space is therefore reduced to that of the formal *natural minimal basis* (NMB), spanning core and valence-shell NAOs only, whereas the residual *natural Rydberg basis* (NRB) of extra-valence NAOs plays practically no significant role in NBO analysis. This condensation of occupancy into the much smaller set of NMB orbitals (allowing the large residual NRB set from the original basis to be effectively ignored) is one of most dramatic and characteristic simplifying features of "natural" analysis.[?] The way to obtain the NAOs is as follows:

1. Each atomic block of the density matrix is diagonalized to obtain a set of non-orthogonal NAOs, namely *pre-NAOs*.

2. An occupancy-weighted orthogonalization is carry out over the highly occupied pre-NAOs of one atom with respect to the other atoms highly occupied pre-NAOs.

3. The negligible occupied pre-NAOs (such Rydberg-type NAOs) of each centre are orthogonalized with the highly occupied pre-NAOs of the same centre through a Gram-Schmidt orthogonalization algorithm.

4. The negligible occupied pre-NAOs of a specific centre are orthogonalized with the negligible occupied pre-NAOs of the other centres by means of a occupancy-weighted orthonormalization.

After carry on this algorithm we obtain a set of orthogonal NAO. The diagonal matrix elements of the density matrix on this basis are the orbitals occupancies. If we sum up the population of all the orbitals belonging to the A atom we obtain the charge of the A atom. Some of the advantages of this populations analysis compare to the Mulliken analysis are that: due to the fact that the NAOs are defined starting from the density matrix, their occupancies will always be between 0 and 2. Another advantage is that as the basis set is extended they converge to well defined values.

Once the density matrix has been converted into the NAOs basis, the bond between atoms arises from the extra-diagonal blocks of the matrix. To determine these *Natural Bond Orbitals* the following steps are followed:

- The NAOs belonging to an atomic block of the density matrix which an occupancy number very close to 2 (> 1.999) are classified as core orbitals and deleted from the density matrix.

- The NAOs belonging to an atomic block of the density matrix which has high occupancy (greater than 2) is identified as lone pair orbitals and they are deleted from the density matrix too.
- Each atomic pair is considered (AB, AC, BC,) and these sub-block two by two are diagonalized (the core and lone pair have been deleted from these sub-blocks). The NBOs are the eigenvalues with high occupancy (normally over 1.90).

If the sum of the occupancy of core, lone pair and NBOs orbitals is lower than the total number of electrons, it is considered that the number of NBOs is not enough and in the next step the occupancy threshold to select the NBOs is lowered. Alternatively, we can search for three centres bonds.

4.1.1 Natural Resonance Theory (NRT)

The *natural resonance theory* (NRT),[?,?] provides an analysis of molecular electron density (correlated or uncorrelated) in terms of resonance structures and weights.
The NRT algorithm is based on representing the one-electron reduced density operator, $\hat{\Gamma}$, as an optimized resonance hybrid of density operators, $\{\hat{\Gamma}_\alpha\}$,

$$\hat{\Gamma} \simeq \sum_\alpha \omega_\alpha \hat{\Gamma}_\alpha \tag{99}$$

where each $\hat{\Gamma}_\alpha$ is the reduced density operator corresponding to an idealized resonance structure wavefunction Ψ_α (determinant of doubly occupied NBOs for the chosen Lewis structure). The NRT resonance weights $\{\omega_\alpha\}$ are constrained to satisfy

$$\omega_\alpha \geq 0, \sum_\alpha \omega_\alpha = 1 \tag{100}$$

in order to preserve the normalization, hermiticity and positivity of the true $\hat{\Gamma}$. NRT analysis can be performed on any wavefunction or density functional method that leads to a one-electron density matrix. Also bond orders and valencies can be evaluated from the final optimized resonance weights $\{\omega_\alpha\}$.

4.2 Atoms In Molecules (AIM)

This method is based in the topological analysis of the charge density function[?] that allows to define the concepts of atom, bond, molecular structure

and structural stability.

The charge density is given by:

$$\rho(\vec{r}) = \int |\psi|^2 d\tau \tag{101}$$

where the square of the wavefunction is integrated over the coordinates of all the electrons except for one. The electronic density is obtained for a fixed nuclear configuration so it is a one-electron function that depends on the nuclear coordinates parametrically.

4.2.1 Topological properties of the charge density

Each topological feature of the electronic density, $\rho(\vec{r})$, whether it be a maximum, a minimum or a saddle point, has associated with it a point in space called *critical point* (cp), where the first derivatives of $\rho(\vec{r})$ vanish:

$$\nabla\rho(\vec{r}) = 0 \tag{102}$$

the diagonalization of the Hessian matrix of the charge density give us the three main curvatures for the electronic density, that would be positive in a minimum or negatives for a maximum. The *rank* of a critical point, denoted by ω is equal to the number of non-zero curvatures of ρ at the critical point. The *signature*, denoted by σ, is the algebraic sum of the signs of the curvatures. The critical point is labelled by giving the duo of values (ω, σ).

With relatively few exceptions, the critical points of charge distribution for molecules at or in the neighbourhood of energetically stable geometrical configurations of the nuclei are all of rank three ($\omega = 3$). It is in terms of the properties of critical points with rank three that the elements of molecular structure are defined. A critical point with $\omega < 3$ is said to be degenerated, Such a critical point is unstable in the sense that a small change in the charge density, as caused by displacement of the nuclei, causes it to either vanish or to bifurcate into a number of stable ($\omega = 3$) critical points.There are just four possible signature values for critical points of rank three.

- (3, -3) All curvatures are negative and ρ is a local maximum at r_c. These points are associated with nuclear positions.
- (3, -1) Two curvatures are negative and ρ is a maximum at r_c in the plane defined by their corresponding axes. ρ is a minimum at r_c along

the third axis which is perpendicular to this plane. These critical points are found between every pair of nuclei which are considered to be linked by a chemical bond and therefore are called *bond critical points*.

- (3, +1) Two curvatures are positive and ρ is a minimum at r_c. in the plane defined by their corresponding axes. ρ is a maximum at r_c along the third axis which is perpendicular to this plane. These critical points are related with the structural concept of ring and thus are called *ring critical points*.

- (3, +3) All curvatures are positive and ρ is a local minimum at r_c. These critical points are called *cage critical points*.

The link between the topological features of the electronic density and the different elements of the molecular structure is based on the gradient vector field of the charge density theory.

4.2.2 Gradient vector field of the charge density

The gradient vector field of the charge density is represented through a display of the trajectories traced out by the vector $\vec{\nabla}\rho$. A trajectory of $\vec{\nabla}\rho$, starting at some arbitrary point, is obtained by calculating $\vec{\nabla}\rho(\vec{r}_0)$, moving a distance $\Delta\vec{r}$ away from this point in the direction indicated by the vector $\vec{\nabla}\rho(\vec{r}_0)$ and repeating this procedure until the path so generated terminates. Some general properties are : The vector $\vec{\nabla}\rho(\vec{r})$ is tangent to its trajectory at each point $\vec{r}$. Every trajectory must originate or terminate at a point where $\vec{\nabla}\rho(\vec{r})$ vanishes, i.e. at a critical point in ρ.

4.2.3 Elements of molecular structure

A (3,-3) critical point, such as occurs at each of the nuclear positions, serves as the terminus, of all the paths starting from and contained in some neighbourhood of the critical point. Thus, the nuclei act as attractors of the gradient vector field of $\rho(\vec{r};\vec{X})$. The result of this identification is that the space of a molecular charge distribution, real space, is partitioned into disjoint regions, the basins, each of which contains one point attractor or nucleus. An atom, free or bound, is defined as the union of an attractor and its associated basin.

Alternatively, an atom can be defined in terms of its boundary. For an atom in a molecule the atomic basin is separated from neighbouring atoms

by interatomic surfaces. The existence of an interatomic surface S_{AB} denotes the presence of a (3, -1) critical point between neighbouring nuclei A and B. Thus, the presence of these surfaces provides the boundaries between the basins of neighbouring atoms. The trajectories which terminate at a (3, -1) critical point define a surface, the interatomic surface S_{AB}. Therefore, the atomic surface S_A of atom A is defined as the boundary of its basin. Generally, this boundary comprises the union of a number of interatomic surfaces, separating two neighbouring basins, and some portions that may be infinitely distant from the attractor.

At the (3, -1) bond critical points a series of pairs of gradient path are originated and they terminate at the neighbouring attractors. Each such pair of trajectories is defined by the eigenvector associated with the unique positive eigenvalue of a (3, -1) critical point. These two unique gradient paths define a line through the charge distribution linking the neighbouring nuclei along which $\rho(\vec{r})$ is a maximum with respect to any neighbouring line. Such a line is found between every pair of nuclei whose atomic basins share common interatomic surface. This line of maximum charge density linking the nuclei is called a *bond path* and the (3, -1) critical point referred to as a *bond critical point* as we mentioned previously.
For a given configuration X of the nuclei, a *molecular graph* is defined as the union of the closures of the bond paths or atomic interaction lines. Pictorically, the molecular graph is the network of bond paths linking pairs of neighbouring nuclear attractors. The network of bond paths thus obtained is found to coincide with the network generated by linking together those pairs of atoms, which are assumed to be bonded to one another on the basis of chemical considerations.

The remaining critical points of rank three occur as consequences of particular geometrical rearrangements of bond paths and they define the remaining elements of molecular structure rings and cages. If the bond paths are linked so as to form a ring of bonded atoms then, a (3, +1) critical point is found in the interior of the ring. The eigenvectors associated with the two positive eigenvalues of the Hessian matrix of ρ at this critical point generate an infinite set of gradient paths which originate at the critical point and define a surface, called the ring surface.

If the bond paths are so arranged s to enclose the interior of a molecule with ring surfaces, then a (3, +3) or cage critical point is found in the interior of the resulting cage. The charge density is a local minimum at a cage critical point. Trajectories only originated at such a critical point and terminate at

nuclei, and at bond and ring critical points.
The number and type of critical point which can coexist in a system with a finite number of nuclei are governed by the Poincaré-Hopf relationship.

$$nb + rc = 1 \tag{103}$$

where n is the number of nuclei, b is the number of bond paths (or atomic interaction lines), r is the number of rings, and c is the number of cages. The collection of numbers (n, b, r ,c) is called the *characteristic set* of the molecule.

4.2.4 Bonds and structure

The value of the charge density at the bond critical point can be used to define a bond order. Multiple bonds do not appear as such in the topology of the charge density. Instead, one finds that the extent of charge accumulation between the nuclei increases with the assumed number of electron pair bonds and this increase is faithfully monitored by the value of ρ at the BCP, ρ_b. Thus, the values of ρ_b obtained for ethane, benzene, ethylene and acetylene can be related through a mathematical equation to the bond orders of 1.0, 1.6, 2.0 and 3.0 respectively. Also, the value of ρ_b is found to increase as a bond length decreases.

The bond path has not necessarily to match up with the internuclear axis, been the former larger. This happens when the bond path is curved, a typical situation when there are strained bonds as in cyclic small hydrocarbons. The differences between the both path angle α_b, the limiting value of the angle subtended at a nucleus by two bond paths, and the geometrical so-called bond angle α_e is important to quantifying the previously mentioned concept of strain in these kind of molecules.

4.2.5 Properties of the Laplacian of the charge density

The Laplacian of ρ, $\nabla^2\rho$, is defined as the trace of the Hessian matrix. It determines where the field is locally concentrated ($\nabla^2\rho(\vec{r}) < 0$) and depleted ($\nabla^2\rho(\vec{r}) > 0$) . Thus, it follows that the value of $\rho(\vec{r})$ is greater than the average of its values over an infinitesimal sphere centred on $\vec{r}$ when the sum of the three curvatures of ρ is negative, that is, when $\nabla^2\rho(\vec{r}) < 0$, and $\rho(\vec{r})$ is less than this average when ($\nabla^2\rho(\vec{r}) > 0$).

The electronic density is a maximum within the interatomic surface associated to the BCP, where it reaches its maximum value. At the same time, the electronic density has a minimum in this point through the bond path. Thus, the generation of an interatomic surface and of an interatomic interaction along the molecular bond is the result of two concomitant effects: the perpendicular contraction of the electronic density that triggers the concentration of electronic charge along the bond path, and the parallel expansion of the electronic density that leads to the depletion of the electronic charge within the surface and the concentration in the basins of the neighbouring atoms. The sign of $(\nabla^2\rho(\vec{r}))$ will determine which of these two effects would dominate.

In the case of covalent bonds in which the electronic density is concentrated at the internuclear region as a result of the distribution of the electrons between the two nuclei, the BCP would have associated a negative value for the Laplacian. On the other hand, when there are interactions between closed-shell systems as ionic bonds, hydrogen bonds, van der Walls complexes, the electronic density on the proximities of the interatomic surface is depleted and it is concentrated in the basins of the neighbouring atoms. Therefore, the BCP has a positive associated Laplacian value.

Using the expression of the virial theorem as a function of the Laplacian is also possible to determine which kind of interaction is taken place between an atom pair.

$$\frac{\hbar^2}{4m}\nabla^2\rho(\vec{r}) = 2G()\vec{r} + V(\vec{r}) \tag{104}$$

where $G(\vec{r})$ represents a kinetic energy having always positive values. $V(\vec{r})$ stands for the potential energy density, and is always negative. Depending on the Laplacian sign one of these two contributions would dominate over the other. In the regions where the Laplacian has a negative value, the stabilization is reached through lowering the potential energy as a result of charge concentration in between the two nuclei. If, in turn, the Laplacian is positive, the kinetic contribution predominates over the potential one and the charge accumulation on this point destabilizes the system. Consequently, in the covalent interactions the bond regions stabilize the system and the Laplacian is negative whereas in the electrostatic-type interactions the Laplacian has positives values in the region between the nuclei.

Thus, the Laplacian sign is a criterion that can be used to determine the ionic/covalent character of the interactions in most of the systems. Never-

theless, there are some exceptions in which this criterion fails as in molecules with weak covalent bonds such a F_2, HOOH, . In these cases, it is used to resort to the energy density in order to classify the interactions.

4.2.6 Energy density

The energy density, $H(\vec{r})$, is defined as :

$$H(\vec{r}) = G(\vec{r}) + V(\vec{r}) \tag{105}$$

where $G(\vec{r})$ and $V(\vec{r})$ are the same kinetic and potential density as described in the equation (104). Using the viral theorem eq. (104) we can related the Laplacian and the energy density as follows:

$$H(\vec{r}) = \frac{1}{4}\nabla^2\rho(\vec{r}) - G(\vec{r}) \tag{106}$$

We can use the sign of the energy density instead of the Laplacian to determine the ionic/covalent character of an interaction. The advantage of this criterion?is that it makes disappear the 2:1 factor existing in equation (104) and which give rise to troubles in the cases where $2G(\vec{r}) > |V(\vec{r})| > G(\vec{r})$. In the case of equation (105), the sign of $H(\vec{r})$ exactly determines which of both contributions, $G(\vec{r})$ or $V(\vec{r})$ dominates.

4.3 Electron localization function

The concepts of electron localization and delocalization are highly relevant in chemistry. Electron localization is essential for descriptive chemistry because in this field one needs to know where local groups of electrons such as core or valence electrons, electron pairs, bonding pairs, unpaired electrons, or π-electron subsystems are placed. The electron localization function (ELF) is a function introduced by Becke and Edgecombe as a mean to measure the electron localization in molecular systems.[27] They showed that the leading term in the Taylor expansion of the spherically averaged conditional pair probability eq.(107)

$$P^{\sigma\sigma}_{cond}(\mathbf{r}, s) = \frac{1}{3}\left[\tau_\sigma - \frac{1}{4}\frac{(\nabla\rho_\sigma)^2}{\rho_\sigma}\right] s^2 + \cdots , \tag{107}$$

where the argument $(\mathbf{r}, s)$ denote spherical average on a shell of radius s about the reference point $\mathbf{r}$, and τ_σ is the positive-definite kinetic energy

density defined by eq. (108)

$$\tau_\sigma = \sum_i^\sigma |\nabla\Psi|^2 \tag{108}$$

and where the contents of the squared brackets in eq. (107) are evaluated at the reference point. Eq. (107) succinctly conveys electron localization information. The smaller the probability of finding a second like-spin electron near the reference point, the more highly localized is the reference electron. Hence, the electron localization is related to the smallness of the expression

$$D_\sigma = \tau_\sigma - \frac{1}{4}\frac{(\nabla\rho_\sigma)^2}{\rho_\sigma} \tag{109}$$

The drawback of eq. (109) as a measure of electron localization is that high localizability is implied by small values of D_σ. Also, D_σ is not bounded from above. Therefore, Becke and Edgecombe proposed an alternative "electron localization function" (ELF) with more desirable features:

$$ELF = (1 + \chi_\sigma^2)^{-1} \tag{110}$$

where

$$\chi_\sigma = D_\sigma / D_\sigma^0 \tag{111}$$

and

$$D_\sigma^0 = \frac{3}{5}(6\pi^2)^{2/3}\rho_\sigma^{5/3}; \tag{112}$$

where D_σ^0 corresponds to a uniform electron gas with spin-density equal to the local value of $\rho_\sigma(\mathbf{r})$. The ratio χ_σ is thus a dimensionless localization index calibrated with respect to the uniform-density electron gas as reference. The transformation of eq. (110) is designed to restrict the ELF's possible values to the range $0 \leq \text{ELF} \leq 1$ with the upper limit ELF corresponding to perfect localization and the value ELF = 0.5 corresponding to electron-gas-like pair probability. Two years later Silvi and Savin[28] showed that ELF measures the excess of kinetic energy density due to the Pauli repulsion in comparison with an uniform electron gas. In the region of space where the Pauli repulsion is strong (single electron or opposite spin-pair behaviour) ELF is close to 1, whereas where the probability of finding the same-spin electrons close together is high, ELF tends to 0. This orbital-independent (*i.e*, with respect to unitary transformations) localization function nicely reveals the location

of atomic shells, and core, binding, and lone electron pairs in atomic and molecular systems.

As in QTAIM analysis, it is possible to partition the molecular space into subsystems (basins) of attractors which allow the calculation of several properties by integration over these basins. The gradien vector field of ELF, $\nabla\eta$, enables one to divide the Euclidian space in basins of attractors where electron pairs are located. These basins are either core basins surrounding a nucleus or valence basins that do not include a nucleus (except for protonated valence basins that include a proton). The number of connections of a given valence basin with core basins is called the synaptic order. A disynaptic valence basin corresponds to a two-center bond, whereas a monosynaptic one characterizes a lone pair. Multicenter bonds, such as three-center two-electron (3c–2e) bonds, are accounted for by polysynaptic basins.[111]

Basin related properties are calculated by integrating a certain property over the volume of the basins. Let N be the number of electrons and $\rho(\mathbf{r}) = \sum_{j=1}^{N} |\phi_j(\mathbf{r})|^2$. For a basin labeled Ω_A, one can define its average population and pair populations by integrating the electron density and the pair densitites as

$$\begin{aligned}
\overline{N}(\Omega_A) &= \int_{\Omega_A} \rho(\mathbf{r})\,\mathrm{d}\mathbf{r} \\
\overline{N}^{\alpha\alpha}(\Omega_A) &= \iint_{\Omega_A} \Gamma^{\alpha\alpha}(\mathbf{r}_1,\mathbf{r}_2)\,\mathrm{d}\mathbf{r}_1\,\mathrm{d}\mathbf{r}_2 \\
\overline{N}^{\beta\beta}(\Omega_A) &= \iint_{\Omega_A} \Gamma^{\beta\beta}(\mathbf{r}_1,\mathbf{r}_2)\,\mathrm{d}\mathbf{r}_1\,\mathrm{d}\mathbf{r}_2 \\
\overline{N}^{\alpha\beta}(\Omega_A) &= \iint_{\Omega_A} \Gamma^{\alpha\beta}(\mathbf{r}_1,\mathbf{r}_2)\,\mathrm{d}\mathbf{r}_1\,\mathrm{d}\mathbf{r}_2
\end{aligned} \tag{113}$$

where the subscript A on Ω indicates that the integration has to be carried out only through the space corresponding to the atomic basin of atom A and $\Gamma^{\alpha\alpha}(\mathbf{r}_1,\mathbf{r}_2)$, $\Gamma^{\beta\beta}(\mathbf{r}_1,\mathbf{r}_2)$ and $\Gamma^{\alpha\beta}(\mathbf{r}_1,\mathbf{r}_2)$ are the same-spin and opposite-spin components of the pair density.[112] Summation of all the atomic populations in a molecule yields the total number of electrons, N.

Multivariate analysis of electron densities. The multivariate analysis is a basic statistical method enabling one to reveal the correlation between

different groups of data. It relies upon the construction of the covariance matrix elements defined by

$$\langle \mathrm{cov}(A,B)\rangle = \langle AB\rangle - \langle A\rangle\langle B\rangle \tag{114}$$

where $\langle A\rangle$, $\langle B\rangle$, and $\langle AB\rangle$ are the averages of the data values and of their product. The diagonal elements of the covariance matrix are the variances

$$\sigma^2(A) = \langle A^2\rangle - \langle A\rangle^2 \tag{115}$$

which measure the dispersion of the data among the group. The square root of the variance is the standard deviation. Finally, the correlation coefficients are the ratios of the covariance matrix elements by the corresponding standard deviation, that is, $[\mathrm{cov}(A,B)]/[\sigma(A)\sigma(B)]$. Positive and negative values fo teh corrrelation coefficients indicate that the A and B data are respectively correlated or anticorrelated, whereas a value close to 0 corresponds to independent behaviours. Applied to electrons distributed among a collectionof adjacent regions spanning the geometrical space occupied by a molecule, the multivariate analysis provides a convenient tool to study electron delocalization. Consider such a partition in M basins for an N electron system. The number of electrons within each region is a quantum mechanical observable to which corresponds the population operator introduced by Diner and Claveries:[113]

$$\hat{N}(\Omega_A) = \sum_i^N \hat{y}(\mathbf{r}_i) \quad \text{with} \quad \hat{y}(\mathbf{r}_i) = \begin{cases} \hat{y}(\mathbf{r}_i) = 1 & \mathbf{r}_i \in \Omega_A \\ \hat{y}(\mathbf{r}_i) = 0 & \mathbf{r}_i \notin \Omega_A \end{cases} \tag{116}$$

In an N-electron system the population operators obey the closure relation[114]

$$\sum_A \hat{N}(\Omega_A) = N \tag{117}$$

The eigenvalues of the population of the population operators, the electron numbers $N(\Omega_A)$, are integers in the range 0, ..., N. As they also obey the closure relationship, the electron count in a region is not independent of those in the other regions and, therefore, theses eigenvalues must be determined simulatneously. The clusure relationship of the basin population operators enables one to carry out the multivariate statistical analysis of the basin populations through the definition of the covariance operator.[114] The expectation values of this operator

$$\langle \mathrm{cov}(\Omega_A,\Omega_B)\rangle = \int_{\Omega_A}\int_{\Omega_B} \Gamma(\mathbf{r}_1,\mathbf{r}_2)\,\mathrm{d}\mathbf{r}_1\,\mathrm{d}\mathbf{r}_2 - \overline{N}(\Omega_A)\overline{N}(\Omega_B) \tag{118}$$

where $\Gamma(\mathbf{r}_1, \mathbf{r}_2)$ denotes the spinless pair density,[112] provide pieces of information about electron delocalization. In particular, the diagonal elements, the variances

$$\sigma^2(\Omega_A) = \int_{\Omega_A}\int_{\Omega_A} \Gamma(\mathbf{r}_1, \mathbf{r}_2)\, \mathrm{d}\mathbf{r_1}\, \mathrm{d}\mathbf{r}_2 - \left[\overline{N}(\Omega_A)\right]^2 + \overline{N}(\Omega_A) \tag{119}$$

are a measure of the quantum mechanical uncertainty of the basin's population, namely, the degree of fluctuation of the electron pair (*i.e.*, the square of the standard deviation), which can be interpreted as the dispersion of the electronic structures. According the partition scheme, the multivariate analysis enables one to build a phenomenological classical model of the charge distribution of a molecule in terms of the superposition of mesomeric structures.

The variance, $\sigma^2(\Omega_A)$, can also be spread in terms of contribution from other basins, the covariance, $\mathrm{cov}(\Omega_A, \Omega_B)$, which has a clear relationship with the so-called delocalization index (DI), $\delta(\Omega_A, \Omega_B)$, defined by Fradera in the AIM framework:

$$\begin{aligned}
\mathrm{cov}(\Omega_A, \Omega_B) &= \langle N(\Omega_A) \cdot N(\Omega_B)\rangle - \langle N(\Omega_A)\rangle\langle N(\Omega_B)\rangle \\
&= \int_{\Omega_A}\int_{\Omega_B} (\Gamma(\mathbf{r}_1, \mathbf{r}_2) - \rho(\mathbf{r}_1)\rho(\mathbf{r}_2))\, \mathrm{d}\mathbf{r}_1\, \mathrm{d}\mathbf{r}_2 \\
&= \int_{\Omega_A}\int_{\Omega_B} \Gamma_{XC}(\mathbf{r}_1, \mathbf{r}_2)\, \mathrm{d}\mathbf{r}_1\, \mathrm{d}\mathbf{r}_2 = -\frac{\delta(\Omega_A, \Omega_B)}{2}
\end{aligned} \tag{120}$$

The DI, $\delta(\Omega_A, \Omega_B)$, accounts for the electrons delocalized or shared between basins Ω_A and Ω_B.[115] As the total variance in a certain basin can be written in terms of covariance, we have

$$\sigma^2(\Omega_A) = -\sum_{B\neq A} \mathrm{cov}(\Omega_A, \Omega_B) = \sum_{B\neq A} \frac{\delta(\Omega_A, \Omega_B)}{2} \tag{121}$$

Form the quantity above one can do the usual contribution analysis (CA), given usually as a percentage:

$$CA(\Omega_A|\Omega_B) = \frac{\mathrm{cov}(\Omega_A, \Omega_B)}{\sum\limits_{A\neq B} \mathrm{cov}(\Omega_A, \Omega_B)} \times 100 = -\frac{\mathrm{cov}(\Omega_A, \Omega_B)}{\sigma^2(\Omega_B)} \times 100 \tag{122}$$

The CA gives us the amin contribution arising from other basins to the variance, that is, the delocalized electrons of basin Ω_B on basin Ω_A, providing a measure of electron pair sharing between two regions of the molecular space.

5 RRKM THEORY

The Rice-Ramsperger-Kassel-Marcus (RRKM) theory is a statistical theory of chemical reactivity that allows to compute unimolecular reaction rates by using characteristics of the potential energy surface.[116] It was developed by Rice and Ramsperger in 1927[117] and Kassel in 1928[118] (RRK theory). In 1952 and taking into account the transition state theory (TST) developed by Eyring in 1935, Marcus generalized this theory to the so called RRKM theory.[119]

5.1 What is a unimolecular reaction?

A unimolecular reaction is defined as any system that evolves in time as a result of some prior stimulus or excitation step. Thus, both dissociation and isomerization are examples of unimolecular processes. Phenomenologically, the kinetics of a unimolecular reaction, A $\rightarrow$ products, is written as

$$-\frac{d\mathbb{N}_A}{dt} = k\mathbb{N}_A \tag{123}$$

which, when integrated, gives rise to the time dependence of the concentration of $A(t)$:

$$\mathbb{N}_A(t) = \mathbb{N}_A(t=0)e^{-kt} \tag{124}$$

where k is the unimolecular rate constant with units of reciprocal time, and $\mathbb{N}_A(t=0)$ is the nuber of molecules of species A at time $t=0$. The rate constant k depends on the internal energy of A, or in the case of an equilibrium ensemble of A, its temperature. For a given state of excitation, the exponential decay is a result of the assumption that the rate is a function only of the concentration of A.

The important questions in the study of unimolecular reactions are: a)what is the initial state produced in the excitation step, b)how fast does the system evolve toward products, c)what are the reaction products, and d) what are the product energy states?.

It was recognized that a dissociation rate depends on the internal energy of the molecule.[120] Thus, all detailed statistical theories of unimolecular reactions begin with the calculation of $k(E)$, the rate constant as a function of the internal energy, E.

The connection between $k(T)$, often called the canonical rate constant, and $k(E)$, the microcanonical rate constant, involves averaging $k(E)$ over the energy distribution

$$k(T) = \int_{E_0}^{\infty} P(E,T)k(E)\,dE \tag{125}$$

where E_0 is the activation energy and $P(E,T)$ is the distribution of internal energies at a given temperature, T. In may aplications the quantity of interest is $k(T)$ however we will work with $k(E)$ since we assume a microcanonical ensemble, i. e., $\mathrm{E}^{TOT} = \text{const}$. It is also worth to stress that the theory can be tested adequately only by comparing the measured and calculated $k(E)$. Once this has been accomplished, the rate constant of any system with a known distribution of internal states can be calculated.

5.2 Introductory remarks

Before going more into details of the RRKM theory, few inroductory remarks about some concepts used for its derviation will be introduced.

Phase space. Statistics involves the counting of states, and the state of a classical particle is completely specified by the measurement of its position **q** and momentum **p**. It is often convenient in statistics to imagine a 2N-dimensional space composed of the N position and N momentum coordinates. Such space is conventionally called "phase space". The counting tasks can then be visualized in a geometrical framework where each point in phase space corresponds to a particular position and momentum. That is, each point in phase space represents a unique state of the particle. In Fig. 43 is shown an example of a phase space plot for a simple pendulum. At point **A** the pendulum is in the maximum positive distance from the bob's neutral point but it's velocity is zero. This is shown as point **A** on the phase space diagram. At **B** the distance of the bob from its neutral poisiton is zero, but its velocity is at a maximum (in a negative sense). The other points of the phase space plot show the relation betwen the velocity and position for other pendulum poisitions. The state of a system of particles corresponds to a certain distribution of points in phase space.

Sum of states. The counting of the number of states available to a particle amounts to determining the available volume in phase space. Let us consider

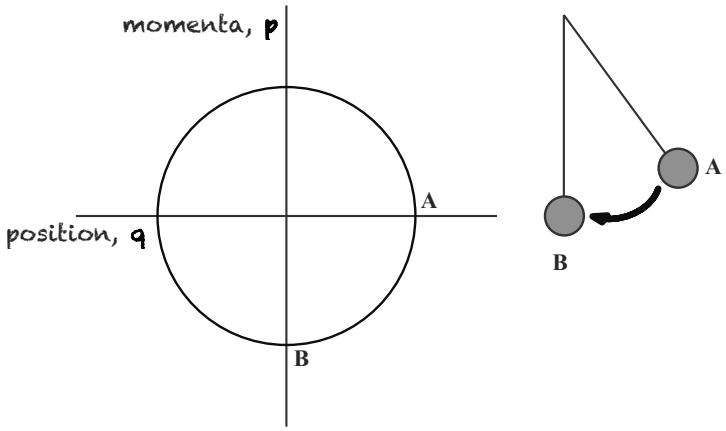


Figure 43: Illustration of phase space for a pendulum.

a molecule consisting of N atoms with a Hamiltonian $H(\mathbf{p}, \mathbf{q})$. The momenta, $\mathbf{p}$ and the position, $\mathbf{q}$, vectors will consist of $m = 3N - 6$ terms, considering that there is no rotational-vibrational coupling. The classical phase space volume of such a system with a maximum energy E is defined by the integral

$$\text{Phase Space Volume} = \int_{H=0}^{H=E} \cdots \int dp_1 \cdots dp_m \, dq_1 \cdots dq_m \tag{126}$$

This volume integral in $2m$-dimensional space, has units of [Joules·sec]^m. In one dimension, the units of phase space are Joule·sec. This is often referred to as a unit of action. According to the uncertainty principle, momentum and position are conjugate quantities which cannot be simultaneously and precisely known, that is $\Delta p \Delta q \geq \hbar/2$. Hence, the smallest allowable unit in phase space must be on the order of h, so that the quantum phase space is divided up into units of h. We can convert the phase space volume into a sum of states simply by dividing by h^m, one unit of h each dimension. The sum of states, $N(E)$, is then

$$N(E) = \frac{1}{h^m} \int_{H=0}^{H=E} \cdots \int dp_1 \cdots dp_m \, dq_1 \cdots dq_m \tag{127}$$

It represents the total number of states for a system corresponding to an energy less than or equal to a specified value, E. Even though the phase space has been divided into quanta of action, this is still considered the classical sum of states because the classical phase space volume is first calculated and converted into quantum states only at the end.

Density of states. A quantity related to the sum of states is the density of states, $\rho(E)$, which is defined as the number of states per unit energy. The number of states in the range E and $E + dE$ is denoted by $W(E)$. It is obtained by integrating d**p**d**q** between $H = E$ and $H = E + dE$, and dividing the resulting volume by h^m:

$$W(E) = \frac{1}{h^m} \int_{H=E}^{H=E+dE} \int dp_1 \cdots dp_m \, dq_1 \cdots dq_m \tag{128}$$

Equation 128 can be expressed as a product of a surface integral evaluated at $H = E$ and the energy interval dE. Since $\rho(E) = W(E)/\, dE$, the density of states is the surface integral

$$\rho(E) = \frac{1}{h^m} \int_{H=E} \cdots \int dp_1 \cdots dp_m \, dq_1 \cdots dq_m \tag{129}$$

The density of states can also be obtained from $N(E)$ by taking the derivative with respect to the energy E, $\rho(E) = dN(E)/\, dE$.

5.3 Derivation of the RRKM statistical theory.

There are several ways to derive the RRKM equation. The one adopted here is based on **classical transition state theory** and was first proposed by Wigner.[121,122] It is worth pointing out that what we call RRKM has also been termed quasi-equilibrium theory (QET)[123] The basic assumptions of statistical theories is that the rate constant $k(E, J)$ depends only on the total energy E and the total angular momentum J. It is assumed that the rate constant does not depend upon where the energy is initially located and that a microcanonical ensemble is maintained as the molecule dissociates. This is equivalent to assuming that IVR (intramolecular vibrational energy redistribution) is rapid compare to the lifetime with respect to dissociation. That is, vibrations are assumed to be strongly coupled by higher order terms (anharmonicities, Fermi resonances, etc) in the expansion of the potential energy function.

The derivation presented here is a simple derivation of the microcanonical rate constant $k(E)$ in which rotations are ignored and in which the location of the transition state is assumed to be fixed at a saddle point and is thus independent of the energy in the system: tight transition state.

As aforementioned the state of a system of particles corresponds to a certain distribution of points in the phase space. The assumption is made

that the total phase space is populated statistically which means that the population density over the whole surface of the phase space is uniform, in other words, the molecule populates the phase space uniformly throughout its dissociation.

In a microcanonical system (constant energy) the phase space is limited to a surface in which the Hamiltonian $H = E$, see Fig. 44. Consequently, the dimensionality of this hypersurface is reduced to $2m - 1$. If the energy, E, is greater than the dissociation energy, E_0, the molecule has a chance to dissociate and thus to reach a part of the hypersurface which is associated with the critical surface.

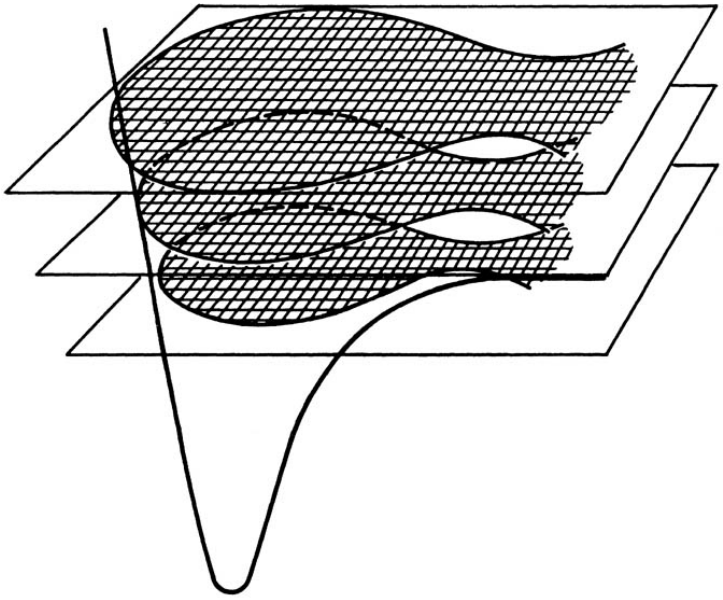


Figure 44: Reaction coordinate with a saddle point. Above it is a diagram of the phase space which varies with the reaction coordinate and with the total energy.

The critical surface is defined as the surface which divides reactants and products. It has a dimension of $2m - 2$. The critical surface is so located that a trajectory, once having passed through it, will proceed on to reaction products without returning back. For reactions with substantial saddle points, the critical surface is located at the saddle point. For systems with small or no saddle points, that is, very small or no reverse activation energy, the location of the critical surface depends upon the total energy, as it would be discussed later.

The rate of reaction is then related to the total flux of reactants which passes through this critical surface.[124] At the critical surface, the molecule is in the process of dissociating along a noe-dimensional reaction coordinate, which is the minimum energy reaction path. It is generally assumed that at the saddle point, the reaction path is perpendicular to all other coordinates, that is, the system is separable. Hence, from now on the spactial and the conjugate momentum $q^\ddagger$ and $p^\ddagger$ (without subscripts) are assigned to these two special coordinates.

Since it was assumed that the phase space is statistically populated, i.e., the population density over the whole surface of the phase space is uniform, the ratio of molecules near the critical surface over the total number of molecules can be expressed as the ratio of the phase space at the dividing surface over the total phase space.

$$\frac{d\mathbb{N}(q^\ddagger, p^\ddagger)}{\mathbb{N}} = \frac{dq^\ddagger\, dp^\ddagger \int \cdots \int_{H=E-\epsilon_t-E_0} \int\, dq_1^\ddagger \cdots dq_{n-1}^\ddagger\, dp_1^\ddagger \cdots dp_{n-1}^\ddagger}{\int_{H=E} \int\, dq_1 \cdots dq_n\, dp_1 \cdots dp_n} \tag{130}$$

where E_0 is the activation energy and ϵ_t is the translational energy associated with the momentum $p^\ddagger$ in the reaction coordinate. Both, the activation energy, E_0, and the translational energy, ϵ_t, must be subtracted from the total energy at the saddle point because these energies are not available for the $n-1$ momenta, $p_i^\ddagger$, and $n-1$ coordinates, $q_i^\ddagger$.

Eq. (130) gives the ratio of molecules whose special coordinates have values that range from $q^\ddagger$ to $q^\ddagger + dq^\ddagger$ and from $p^\ddagger$ to $p^\ddagger + dp^\ddagger$ over the total phase space, at any instant time. The rate of reaction is obtained from the time derivative of the molecules near the critical surface, that represents the flux of molecules passing through the critical region.

$$Flux = \frac{d\mathbb{N}(q^\ddagger,\, dp^\ddagger)}{dt} \tag{131}$$

Because it was assumed that the reaction coordinate is perpendicular to, and separable from all other coordinates the time derivative involves only the $dq^\ddagger\, dp^\ddagger$ term:

$$\frac{d\mathbb{N}(q^\ddagger p^\ddagger)}{dt} = \frac{\mathbb{N}\frac{dq^\ddagger\, dp^\ddagger}{dt} \int_{H=E-\epsilon_t-E_0} \int\, dq_1^\ddagger \cdots dq_{n-1}^\ddagger\, dp_1^\ddagger \cdots dp_{n-1}^\ddagger}{\int_{H=E} \int\, dq_1 \cdots dq_n\, dp_1 \cdots dp_n} \tag{132}$$

noting that

$$\frac{dq^\ddagger}{dt} = \frac{p^\ddagger}{\mu^\ddagger} \Rightarrow \frac{dq^\ddagger\, dp^\ddagger}{dt} = \frac{p^\ddagger\, dp^\ddagger}{\mu^\ddagger} \tag{133}$$

where $\mu^\ddagger$ is the reduced mass of the two separating fragments, Eq. (132) becomes

$$\frac{d\mathbb{N}(q^\ddagger, p^\ddagger)}{dt} = \frac{\mathbb{N}\frac{p^\ddagger\, dp^\ddagger}{\mu^\ddagger} \int_{H=E-\epsilon_t-E_0} \int\, dq_1^\ddagger \cdots dq_{n-1}^\ddagger\, dp_1^\ddagger \cdots dp_{n-1}^\ddagger}{\int_{H=E} \int\, dq_1 \cdots dq_n dp_1 \cdots dp_n} \tag{134}$$

Since the energy in the reaction coordinate is

$$\epsilon_t = \frac{p^{\ddagger 2}}{2\mu^\ddagger} \Rightarrow\, d\epsilon_t = \frac{p^\ddagger\, dp^\ddagger}{\mu^\ddagger} \tag{135}$$

and Eq. (134) can be written as:

$$\frac{d\mathbb{N}(q^\ddagger, p^\ddagger)}{dt} = \frac{\mathbb{N}\, d\epsilon_t^\ddagger \int_{H=E-\epsilon_t-E_0} \int\, dq_1^\ddagger \cdots dq_{n-1}^\ddagger\, dp_1^\ddagger \cdots dp_{n-1}^\ddagger}{\int_{H=E} \int\, dq_1 \cdots dq_n\, dp_1 \cdots dp_n} \tag{136}$$

Eq. (136) expresses the reaction rate (molecules per unit time) in terms of $\mathbb{N}$, the number of molecules, multiplied by the rate constant $k(E, \epsilon_t)$.

$$\frac{d\mathbb{N}(q^\ddagger,\, dp^\ddagger)}{dt} = \mathbb{N}k(E, \epsilon_t) \tag{137}$$

$k(E, \epsilon_t)$ is expressed in terms of a ratio of phase space areas

$$k(E, \epsilon_t) = \frac{d\epsilon^\ddagger \int_{H=E-\epsilon_t-E_0} \int\, dq_1^\ddagger \cdots dq_{n-1}^\ddagger\, dp_1^\ddagger \cdots dp_{n-1}^\ddagger}{\int_{H=E} \int\, dq_1 \cdots dq_n\, dp_1 \cdots dp_n} \tag{138}$$

These phase space areas can be converted into densities of states. In fact, the denominator of Eq. (138) is just the density of states multiplied by the factor h^n. The numerator is an integral over one less dimension, so that it is a density multiplied by h^{n-1}. Thus, we can rewrite the rate constant as:

$$k(E, \epsilon_t) = \frac{\rho(E - E_0 - \epsilon_t)h^{n-1}}{\rho(E)h^n} \Rightarrow k(E, \epsilon_t) = \frac{\rho(E - E_0 - \epsilon_t)}{h\rho(E)} \tag{139}$$

Eq. (139) express the rate constant in terms of the total energy, E, and the translational energy of the departing fragments at the transition state, ϵ_t. As shown in Fig.**??** there are many different ways for the reaction to pass through the transition state region. They differ in how the available energy, $E - E_0$, is partitioned between the internal energy of the transition state and the translational energy of the fragments. Eq. (139) is a state-to-state reaction rate constant. In order to obtain the total dissociation rate we must integrate over all the different translational energies in the transition state.

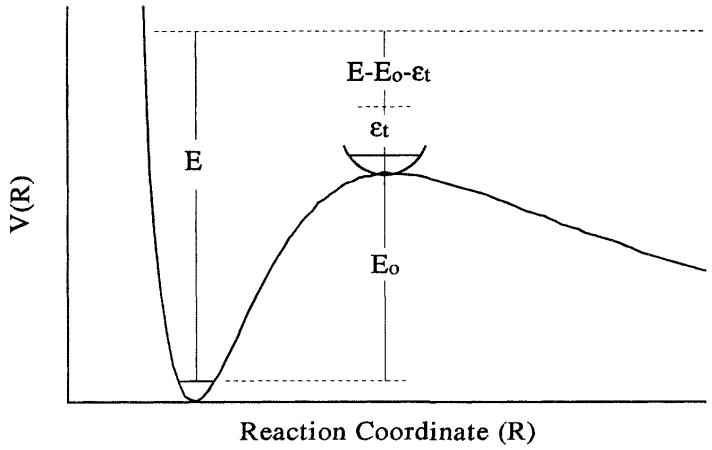


Figure 45: The reaction coordinate and the partitioning of the energy in the transition state $(E - E_0)$ between the translational energy, ϵ_t, and the vibrational energy of the modes normal to the reaction coordinate.

$$k(E) = \frac{\int_0^{E-E_0} \rho^{\ddagger}(E - E_0 - \epsilon_t) d\epsilon_t}{h\rho(E)} = \frac{N^{\ddagger}(E - E_0)}{h\rho(E)} \tag{140}$$

where $N^{\ddagger}(E-E_0)$ is the sum of states at the transition state from 0 to $E-E_0$. Because in the calculation of densities and sums of states, the reactant and transition state symmetries are ignored we must multiplied the rate constant in Eq. (140) by the reaction symmetry, σ,

$$k(E) = \frac{\sigma N^{\ddagger}(E - E_0)}{h\rho(E)} \tag{141}$$

Eq. (141) is known as the RRKM equation.

5.3.1 Assumptions in the RRKM theory.

The RRKM equation has been derived in terms of a classical mechanical flux in phase space. The conversion of this equation into a quantum mechanical theory is done converting the phase space volume and surfaces into sum and density of states by dividing these quantities by h^{n-1}. However the zero point energy (ZPE) must be taken into account somehow. With this purpose all energies are referred at the zero point energy. Hence, E_0 is defined as the

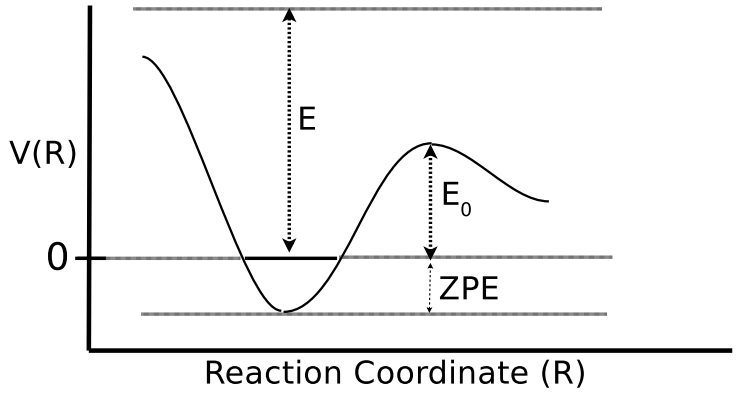


Figure 46: Zero energy reference to compute $k(E)$.

dissociation energy at 0 K (see figure 46).

The fundamental assumption of the statistical theory is that the molecule populates the phase space uniformly throughout its dissociation. A micro-canonical ensemble is maintained. This is only true when the intramolecular vibrational energy redistribution (IVR) is very fast compared to the rate of reaction. That is known as intrinsic RRKM behaviour.

$$\mathbb{N}(t) = \mathbb{N}(0)e^{-k(E)t} \Rightarrow \text{Intrinsic RRKM behaviour} \tag{142}$$

The second assumption is that all the molecules within the region of phase space bounded by $q^\ddagger$ and $q^\ddagger + dq^\ddagger$ and $p^\ddagger$ and $p^\ddagger + dp^\ddagger$ lead to products. That is, all the molecules that cross the critical surface lead to products. There is no recrossings. These recrossings would reduce the rate so that RRKM theory would overestimate the rate constant. The TS is located in the region where the recrossings are minimized.

The third assumption is that the coordinates, $dq^\ddagger$ and $dp^\ddagger$, are perpendicular to all other coordinates and therefore, separable from the other coordinates. This assumption becomes less valid when the energy increases and the coupling between the reaction coordinate and the rest of the modes becomes more important.

5.4 Loose TS

For channels with no reverse activation energy (loose-TS) the TS is more difficult to define because there is no saddle point along the reaction path. In these cases, the TS is located on the basis of the minimum sum of states and, as aforementioned, its location varies with the internal energy.

To compute the RRKM rate constants for the loose-TS we adopted the microcanonical variational transition state theory (μVTST) in its vibrator formulation.[125,126] First we perform a scan along the reaction coordinate. Then, at each point of the scan we do an optimization, freezing this internal coordinate and the energy is computed. The Hessian matrices describing the modes orthogonal to the reaction path are evaluated according to the standard procedure of Hu and Hase,[126] and the sum of states are calculated for teh corresponding optimized strucure. We repeat this procedure for each internal energy considered and the structure corresponding to the minimum sum of states is assigned as the TS. An example of a loose-TS is the formamide neutral loss reaction: $[\mathrm{M(formamide)}]^{2+} \longrightarrow \mathrm{M}^{2+}$ + formamide. In this example, the reaction coordinate corresponds to the Ca–O distance and thus we used this as the scanned coordinate. In Fig47 it can be observed how the location of the TS (minimum sum of states) with respect to the reaction coordinate changes with the internal energy of the reactant molecule, getting closer to the reactants as the molecule's internal energy increases.

5.5 External rotations

RRKM theory treats modes either as active, which exchange energy freely, or adiabatically, which remain in the same quantum number during the unimolecular decomposition. Normally, the modes treated as active are all the normal modes. If the external rotational degrees of freedom are to be taken into account there are different ways of doing so in the framework of RRKM theory.[127] In this study we have considered an *almost symmetric top*, where $I_x \approx I_y$, for which rotational energy is given by the following approximation:[128]

$$E_{rot}(J,K) = \left(\frac{1}{I_x} + \frac{1}{I_y}\right)\frac{\left[J(J+1) - K^2\right]\hbar^2}{4} + \frac{K^2\hbar^2}{2I_z} \tag{143}$$

that can be written as

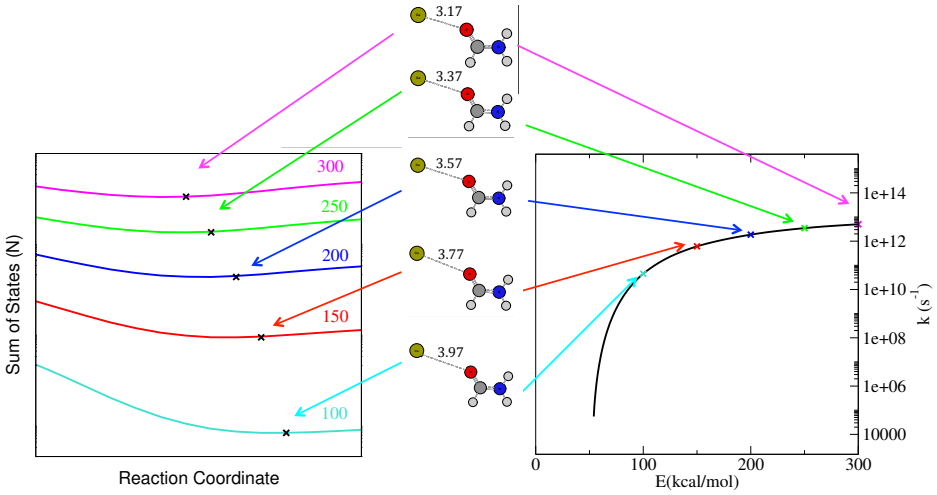


Figure 47: On the left side the sum of states *vs.* the reaction coordinate for formamide–Ca^{2+} at different internal energies (all in kcal mol^{-1}). The crosses mark the minimum of the sum of states for each energy. On the right side the microcanonical rate constant k is represented *vs.* the internal energy. The crosses mark the $k(E)$ values corresponding to the minimum sum of states for specific internal energies. In the middle the formamide–Ca^{2+} geometries are represented for each of the points with the values (in Å) for Ca–O distance.

$$E_{rot}(J,K) = BJ(J+1) + (A-B)K^2 \tag{144}$$

where $A = \hbar^2/2I_z$, $B = \left(\frac{1}{I_x} + \frac{1}{I_y}\right)\frac{\hbar^2}{4}$, $J = 0, 1, 2, \ldots$ and $K = 0, \pm1, \pm2, \ldots, \pm J$. The symmetry axis is the z-axis and the other two moments of inertia are $I_x \approx I_y$. A symmetric top can be visualized as a rotating cylinder. For a given J, the cylinder can rotate in a total of $2J+1$ orientations, each with a different K quantum number which determines its projection along the symmetry axis. Fig.48 shows the case of a prolate top rotating with $K \approx J$ and $K = 0$.

The quantum number J is a constant of motion and therefore is always adiabatic, *i.e.* there is no exchange. On the other hand, the quantum number K can be treated as an active rotor —it allows energy exchange between vibrational and rotational modes– or as an adiabatic rotor.

If we assume that K is conserved, and thus treated as adiabatic, the RRKM unimolecular rate constant for energy E and specific values of J and

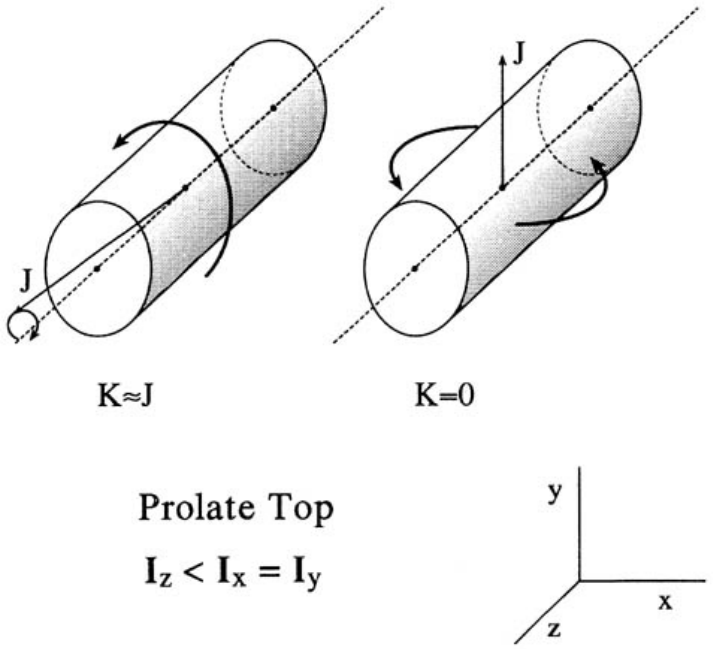


Figure 48: Prolate symmetric top molecule represented as a cylinder. The rotational motion due to J and K are indicated.

K is[124]

$$k(E,J,K) = \frac{\sigma}{h}\frac{N^{\ddagger}\left[E - E_0 - E^{\ddagger}_{rot}(J,K)\right]}{\rho\left[E - E_{rot}(J,K)\right]} \tag{145}$$

where $E^{\ddagger}_{rot}$ is the rotational energy for the transition state and E_{rot} is the same for the minimum energy structure (both obtained by eq.143). Within this assumption (J and K being adiabatic) we consider three particular cases for partitioning the rotational energy within the molecular axes. In case **1**, all the rotational energy is placed on the x,y-plane, and corresponds to $K = 0$, and

$$J = \frac{\left(-1 + \sqrt{1 + 4\frac{E_{rot}}{\overline{B}}}\right)}{2}; \quad \overline{B} = \left(\frac{1}{I_x} + \frac{1}{I_y}\right)\frac{\hbar^2}{4} \tag{146}$$

where E_{rot} is the rotational energy of the ion after collision. In case **2** the rotational energy is equally distributed among the three axes, so $E_{x,y} = 2/3E_{rot}$ and $E_z = 1/3E_{rot}$. Finally, in case **3** all the rotational energy is placed along the z-axis such that

$$K = \sqrt{\frac{1}{3}\frac{E_{rot}}{A - \overline{B}}} \tag{147}$$

The $k(E, J)$ microcanonical rate constant with K active can be determined by two different approaches. One approach for treating the K-dependent term as an active degree of freedom is to calculate the density of states for the reactant and the sum of states for the TS by summing over contributions from all possible values of K to give[127]

$$k(E, J) = \frac{\sigma}{h}\frac{\sum_{K=-J}^{K=J} N^{\ddagger}\left[E - E_0 - E^{\ddagger}_{rot}(J, K)\right]}{\sum_{K=-J}^{K=J} \rho\left[E - E_{rot}(J, K)\right]} \tag{148}$$

In the second approach, the unimolecular rate constant, $k(E, J)$, for a total energy E and angular momentum J is written as[127]

$$k(E, J) = \sigma h \frac{N^{\ddagger}\left[E - E^{\ddagger}_{rot}(J) - E_0\right]}{\rho[E - E_{rot}]} \tag{149}$$

where $N^{\ddagger}$ and ρ are written as convolutions between the densities and the sum of states for the internal degrees of freedom and the active external rotation, respectively

$$N^{\ddagger}(E^{\ddagger}) = \int_0^{E^{\ddagger}} N^{\ddagger}_{vib}(E)\rho_{rot}(E^{\ddagger} - E\, dE \tag{150}$$

$$\rho(E_\nu) = \int_0^{E_0} \rho_{vib}(E)\rho_{rot}(E_\nu - E)\, dE \tag{151}$$

$E^{\ddagger}$ is the active energy of the TS and E_ν is the active energy of the energized reactant; the total energy is $E = E_\nu + E_{rot}(J) = E^{\ddagger} + E^{\ddagger}_{rot}(J) + E_0$ with $E_{rot}(J)$ and $E^{\ddagger}_{rot}(J)$ being the adiabatic rotational energies of the reactant and the TS respectively.

For all the calculations, the sum and density of states were computed using a semiclassical state counting.[129–131] We used the quantum RRKM model which assumes that ZPE flows freely within the molecule.[132,133] All calculations were performed using the RRKM code given by Zhu and Hase.[134]

6 Direct chemical dynamics simulations

Since the early 1960s classical trajectory simulations have been used to study the atomistic dynamics of chemical reactions, energy transfer, and molecular motion.[135,136] The components of a classical trajectory simulation are[137] (1) developing or choosing a potential energy surface for the chemical problem under investigation; (2) selecting initial conditions for the ensemble of trajectories to be calculated; (3) numerical integration of the classical equatins of motion, that is, either Newton's or Hamilton's,[?,138] to determine the atomic-level motion for each trajectory; and (4) transformation of the trajectorie's final atomic coordinates and momenta to properties that may be compared with experiment and/or theoretical model. Included in this last comonent are bond lengths and angles to identify product structures; product vibrational, rotational, and translational energies; quantum numbers for vibrational and rotational degrees of freedom; the amount of energy in individual molecular degrees of freedom; and scattering angles.

During the classical simulation, the motion of the atoms in the chemical system is determined by solving the classical equations of motion, which can be either Newtonian

$$-\frac{\partial V}{\partial q_i} = m_i \frac{\mathrm{d}^2 q_i}{\mathrm{d}t^2} \tag{152}$$

or Hamiltonian

$$\frac{\partial H_n}{\partial p_i} = \frac{\mathrm{d}q_i}{\mathrm{d}t}, \quad \frac{\partial H_n}{\partial q_i} = -\frac{\mathrm{d}p_i}{\mathrm{d}t}, \tag{153}$$

where $H = T(\mathbf{p}) + V(\mathbf{q})$ is the molecular system's total energy, and the index i encompasses all the atoms' coordinates and momenta. We know the kinetic

energy experssion $T(\mathbf{p})$ –essentially, it's $T = p^2/2m$ for one atom. What we need to know for this calculation is the system's potential energy function (or surface), $V(\mathbf{q})$, as well as the gradient of the potential (and in some cases also the Hessian).

Until the 1990s the standard way to perform a classical trajectory chemical dynamics simulation was to represent the potential energy surface, $V(\mathbf{q})$, by an analytic potentail energy function. For example, we can model $V(\mathbf{q})$ for a cluster of rare gas atoms by a sum of Lennard-Jones potentials. The molecular mechanics (MM) model,[139] consisting of functions to represent atomic stretching, bending, wagging, and torsional motions, was developped to describe $V(\mathbf{q})$ for organic and biological molecules with thermal energies. However, this function generally does not correctly describe a PES for some processes as bond-breaking, for which highly excited molecules undergo bond rupture reactions to form new molecules. Another approach, for a small molecular system consisting of a few atoms, is to calculate a high density of points in the coordinate space and fit them with arbitrary functions to give a smooth and accurate PES. However, developing such a function is a formidable and lengthy task that becomes computationallly impractical as the number of atoms of the system becomes large. A different technique to obtain $V(\mathbf{q})$ for large reactive system's is to derive an analytic potential energy function for the degrees of freedom thought to be most critical for the reaction dynamics using electronic structure calulations. Then, empirical analytic potential energy terms are used for the remaining degrees of freedom.[140] All these approaches for representing $V(\mathbf{q})$ have limitations and contain ambiguities and difficulties.

Trajectories on the Fly With increased computer speed and more powerfull computer algorithms, it has become possible to perform *direct dynamics simulations*,[141] also known as "on-the-fly calculation". In a direct dynamics simulation, the classical equations of motion are numerically integrated without the need for an analytic potential energy function.? Instead, the trajectories are integrated on the fly, with the potential energy $V(\mathbf{q})$ and derivative $\partial V/\partial q_i$ obtained directly from an electronic structure theory by solving at each numerical integration step the time-independent Schörredinger equation

$$H_e(\mathbf{r};\mathbf{q})\Psi(\mathbf{r};\mathbf{q}) = E_e(\mathbf{q})\Psi(\mathbf{r};\mathbf{q}) \tag{154}$$

where $\mathbf{r}$ and $\mathbf{q}$ are the electron and nuclear coordinates, and H_e, Ψ, and E_e are the electrons' Hamiltonian operator, wavefunction, and energy at

the specific nuclear configuration given by the set of coordinates $\mathbf{q}$. The potentail energy $V(\mathbf{q})$ is the sum of the electronic energy $E_e(\mathbf{q})$ and the nuclear-nuclear repulsion $V_{NN}(\mathbf{q})$. Eq. (154) can be solved either by means of a wavefunction based method or by means of density functional theory. Since the time-independent Schörredinger equation (154) is solved at each step of the trajectory integration, providing the potential energy $V(\mathbf{q})$ and gradient $\partial V/\partial q_i$, these kind of direct dynamics simulations are called Born-Oppenheimer (BO) direct dynamics.[142,143]

Time-step Over any arbitrary time interval, the relationship between two positions is given by

$$q(t_2) = q(t_1) + \int_{t_1}^{t_2} \frac{p(t)}{m}\,\mathrm{d}t \tag{155}$$

Similarly, the relationship between two momentum vectors is given by

$$p(t_2) = p(t_1) + m\int_{t_1}^{t_2} a(t)\,\mathrm{d}t \tag{156}$$

for the sake of simplicity we have remove the subindex i for atoms coordinate and momenta. It is almost never possible to write down analytical expressions for the position and momentum components of the phase space trajectory as a function of time. However, using Euler's approximation eqs. (155) and (156) can be rewriten as

$$\mathbf{q}(t + \Delta t) = \mathbf{q}(t) + \frac{\mathbf{p}(t)}{m}\Delta t \tag{157}$$

and

$$\mathbf{p}(t + \Delta t) = \mathbf{p}(t) + m\mathbf{a}(t)\Delta t \tag{158}$$

this approximation being exact in the limit of $\Delta t \to 0$. Thus, given a set of intial positions and momenta (initial conditions), and a means for computing the forces acting on each particle at any instant (and thereby deriving the acceleration), we have a formalism for simulating a trajectory. The way of choosing the intial constions will be explained in more details later on. The mean for computing the forces, as aforementioned is the electronic structure theory used to solve the time-independent Schrödinger equation.

While the use of eqs. (157) and (158) seems entirely strightforward, the finite time step introduces very real practical concerns. Fig. 49 illustrates the variation of a single momentum coorrdinate of some arbitrary phase space trajectory, which is described by a smooth curve. When the acceleration is computed for a point on the true curve, it will be a vector tangent to the curve. If the curve is not a straight line, any mass-weighted step along the tangent (eq. (158)) will necessarily result in a point off the true curve. There is no guarantee that computing the acceleration at this new point will lead to a step that ends in the vicinity of the true cuve. Indeed, whit each additional step, it is quite possible that we will move further and further away from the true trajectory. The problem is compoounded for position coordinates, since the velocity vector being used is already only an estimate derived from eq. (158), *i.e.*, there is no guarantee that it will even be tangent to the true curve when a point on the true curve is taken.

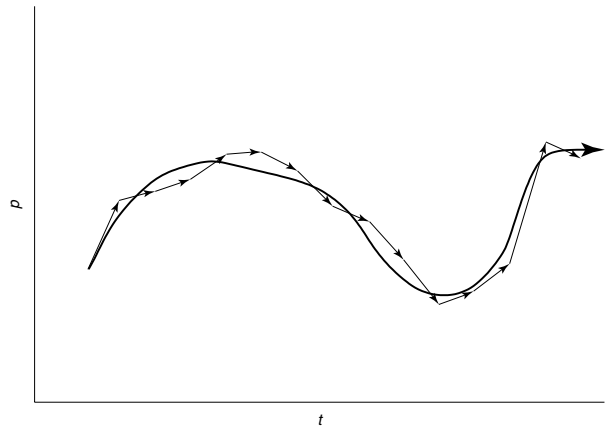


Figure 49: An actual phase-space trajectory (bold curve) and an approximate trajectory generated by repeated application of eq. (158) (series of arrows representing individual time steps). In the illustration, the approximate trajectory hews relatively closely to the actual one, but this will not be the case if too large a time step is used.

In the limit of an infinitesimally small time step, we will recover eq. (155) and (156). But, since each time step requires a computation of all of the molecular forces (at least), which is computationally intensive, we do not want to take too small a time step, or we will not be able to propagate the trajectory for any chemically interesting length of time. What then is the optimal length for a time step that balances numerical stability with chemical utility? The general answer is that it should be at least one and preferably

two orders of magnitude smaller than the fastest periodic motion within the system. In a typical molecular system, the fastest motion is bond vibration which, for a heavy-atom–hydrogen bond has a period of about 10^{-14} s. Thus, for a system containing such bonds, an integration time step Δt should not much exceed 0.1 fs. This rather short time step means that simulations are rarely run for more than some 100 fs of simulation time.

Integration algorithms. Using Euler's approximation and taking integration steps in the direction of the tangent is a particularly simple integration approach, and as such is not particularly stable. Considerably more sophisticated integration schemes have been developed for propagating trajectories. If we restrict ourselves to consideration of the position coordinate, most of these schemes derive from approximate Taylor expansions in $\mathbf{r}$, *i.e.*, making use of

$$\mathbf{q}(t+\Delta t) = \mathbf{q}(t) + \mathbf{v}(t)\Delta t + \frac{1}{2!}\mathbf{a}(t)(\Delta t)^2 + \frac{1}{3!}\frac{d^3\mathbf{q}(\tau)}{dt^3}\bigg|_{\tau=t}(\Delta t)^3 + \cdots \quad (159)$$

where the abbreviations $\mathbf{v}$ and $\mathbf{a}$ for the first and second time derivatives, repectively, of the position vector $\mathbf{q}$ were used.

One such method, first used by Verlet in 1967,[144] considers the sum of the Taylor expansions corresponding to forward and reverse time steps Δt. In that sum, all odd-order derivatives disappear since the odd powers of Δt have opposite sign in the two Taylor expansions. Rearranging terms and truncating at second order (which is equivalent to truncating at third-order, since the third-order term has a coefficient of zero) yields

$$\mathbf{q}(t+\Delta t) = 2\mathbf{q}(t) - \mathbf{q}(t-\Delta t) + \mathbf{a}(t)(\Delta t)^2 \quad (160)$$

Thus, for any particle, each subsequent position is determined by the current position, the previous position, and the particle's acceleration (determined from the forces on the particle). For the very first step (for which no position $\mathbf{q}(t-\Delta t)$ is available) one might use eqs. (157) and (158).

The Verlet scheme propagates the position vector with no reference to the particle velocities. However, often one wants knowledge of the momentum coordinates is also essential. To propagate the position and velocity vectors in a coupled fashion, a modification of Verlet's approach called the leapfrog algorithm has been proposed. An even better implementation of the same basic algorithm is the so-called velocity Verlet[85] algorithm, where

positions, velocities, and accelerations at time $t + \Delta t$ are obtained from the same quantities at time t in the following way:

$$\begin{aligned} \mathbf{q}(t+\Delta t) &= \mathbf{q}(t) + \mathbf{v}(t)\Delta t + (1/2)\mathbf{a}(t)\Delta t^2 \\ \mathbf{v}(t+\Delta t/2) &= \mathbf{v}(t) + (1/2)\mathbf{a}(t)\Delta t \\ \mathbf{a}(t+\Delta t) &= -(1/m)\nabla V(\mathbf{q}(t+\Delta t)) \\ \mathbf{v}(t+\Delta t) &= \mathbf{v}(t\Delta + t/2) + (1/2)\mathbf{a}(t+\Delta t)\Delta t \end{aligned} \tag{161}$$

Initial conditions. The selection of initial conditions will be briefly summarized considering the collision induced dissociation process studied in Part II of this PhD thesis, namely, Ar + formamide-M^{2+} ion. The properties to be sampled in this case are the rotational and vibrational energies of the formamide-M^{2+} ion and teh relative properties of Ar + formamide-M^{2+}. The formamide-M^{2+} molecule may be considered as a symmetric top and its rotational energy may be sampled from a thermal distribution at temperature T_r[145] or be defined as a specific rotational state with quantum numbers J and K. For most polyatomic molecules and for room temperature, the rotational energy distribution is classical and the total angular momentum j and its component on the symmetric top axis j_z are selected by sampling this distribution. Since no specific rotational state is selected in the experiments, we used the classical distribution at a temperature T to sample the rotational energy.

The vibrational energy of the polyatomic formamide-M^{2+} may be that for a specific vibrational state specified by quantum numbers n or chosen for a vibrational temperature T_v. For the latter, the quantum number n_i for each vibrational mode is randomly sampled from the probability distribution

$$P(n_i) \propto \exp\left[-(n_i + \frac{1}{2})h^{\nu_i}/k_B T_v\right] \tag{162}$$

The energy for each individual normal mode is then $E_i = (n_i + 1/2)h\nu_i$. The coordinate Q_i and momenta P_i are chosen by sampling a random phase for teh mode. The $\mathbf{Q}$ and $\mathbf{P}$ vectors are then transformed to Cartesian coordinates and momenta using the normal mode eigenvector.

The relative Ar + formamide-M^{2+} properties include (1) the impact parameter b, which may be chosen randomly between 0 and b_{max}; (2) random rotation of formamide-M^{2+} about its Euler angles; and (3) the Ar + formamide-M^{2+} relative collision energy.

Sampling at a Transition State Once the transition state is identified and the energy transfer obtained, it is possible to use dynamics in order to study how a system evolves from the transition state towards different reaction channels.

Thus, in some cases it is meaningful to initialize trajectories at a TS with either initial conditions for a constant T in accord with TST or at constant energy for a unimolecular reaction in accord with RRKM. This TS sampling allows proper ZPE conditions at the TS and quasiclassical sampling to the TS energy levels. Quasiclassical refer to the selection of the initial conditions in order to represent quantum mechanical vibrational and rotational energy levels. Canonical, constant temperature, sampling at a TS is identical to canonical sampling for a polyatomic molecule, as described above for CID reaction sampling, except there are only 3N-7 vibrational degrees of freedom (nonlinear polyatomic) instead of N-6. The new degree of freedom at the TS is reaction coordinate translation, which is treated clasically and whose probability distribution is

$$P(E_t^\ddagger) = \frac{\exp -E_t^\ddagger / k_b T}{k_B T}, \tag{163}$$

which may be sampled by its cumulative distribution function

$$E_t^\ddagger = -kT \ln(1 - R), \tag{164}$$

where R is a freshly chosen random number in the range oof 0–1. $E_t^\ddagger$ is related to the reaction coordinate momentum via $E_t^\ddagger = (P_t^\ddagger)^2/2$ and the normal mode eigenvector of the TS is used to transform to Cartesian momentum.

For statistical unimolecular decomposition at constant energy, there is a microcanonical ensemble of states for the dissociating molecules and every vibrational/rotational state at the TS with energy $E_{v,r}^\ddagger$ in the range 0 to $E - E_0$ has equal probability of being populated.[146,147] Here E is the energy of the unimolecular reactant above its ZPE level and E_0 is the difference between the TS and reactant ZPE levels; that is, the quantum threshold. The total TS energy $E^\ddagger = E - E_0$ is the sum $E^\ddagger = E_{v,r}^\ddagger + E_t^\ddagger$. With the TS vibrational/rotational state chosen randomly to give $E_{v,r}^\ddagger$, the remaing energy is added to reaction coordinate translation.

When sampling the ZPE level at a TS, either quasiclassical sampling, with random normal mode phases, or Wigner sampling may be used to transform the ZPE level to normal mode coordinates and momenta. These two sampling

algorithms were found to give similar dynamics for trajectories initializated at a TS ZPE level.[148]

CID direct dynamics simulations. For complex problems with many atoms the potential energy surface may be represented as a combination of two components; that is, atoms treated directly by a QM method and the remainder by analytic MM functions. Such a simulation is referred to as QM + MM or QM/MM.[143]

In the direct dynamics simulations performed to molde the CID process between Ar and formamide-M^{2+} (M=Ca, Sr) there is no need to treat explicitly by quantum mechanics the Ar atom, since no bonds will form or break between Ar and the ion. Hence, in order to reduce computation time, the interaction Ar–formamide-M^{2+} is described by an analytical potential, MM. Thus, the potential energy used in the simulations is

$$V = V_{ion} + V_{Ar--ion} \tag{165}$$

where V_{ion} and V_{Ar-ion} are the ion intramolecular and the Ar–formamide-M^{2+} interaction potentials, respectively. A QM description is used to treat the intramolecular ion potential (V_{ion}), while the ion-projectile intermolecular interaction (V_{Ar-ion}) is treated via the analytical potential developed by Meroueh and Hase to simulate CID of protonated peptides.[78] This potential is a sum of two-body terms between Ar and each of the atoms of the molecular ion:

$$V_{Ar--ion} = \sum_i a_i \exp(-b_i r) + \frac{c_i}{r^9} \tag{166}$$

where r is the Ar–ion-atom distance and the a, b and c coefficients are obtained by fitting the analytic potential to the *ab initio* interaction potential. The values for parameters a, b and c for the Ar-formamide and Ar-Ca^{2+} interactions were taken from ref.[78] and ref.[149] Parameters for the Ar-Sr^{2+} interaction were found by fitting eq. (166) to the Ar-Sr^{2+} potential energy curve obtained at the QCISD(T)(full) level, together with a 6-31++G(d,p) basis set expansion for Ar atom and a Stuttgart basis set with pseudo potential for Sr^{2+} cation.[149] BSSE was taken into account using the counterpoise method.[150] The ab initio and fitted curves, as well as a, b and c parameters are shown in Fig. 50. Note that the curve is fitted with a purely repulsive energy function in order to better describe the repulsive wall that is the most important feature in CID for the energies considered here.[78]

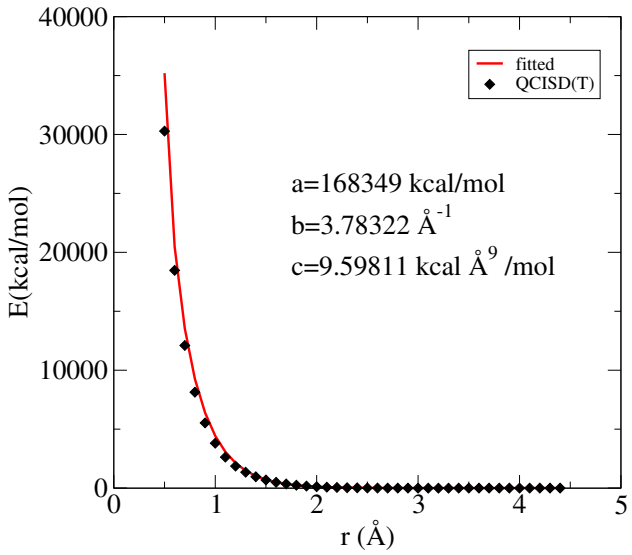


Figure 50: QCISD(T)/6-31++G(d,p) *ab initio* (diamonds) and fitted (solid line) potential energy curves for the Ar-Sr^{2+} interaction.

A general and broadly applicable approach for performing a classical trajectory chemical dynamics simulation is to obtain the potential energy function, and its gradient and Hessian, directly from an electronic structure theory.[142] The latter is a componenet of a quantum chemistry software package such as Gaussian09[84] so that the software needed for such a chemcial dynamics simulation requires to interface the technology of a chemical dynamics simulation with that of quantum chemistry. We used VENUS[80,81] package for the propagation of the trajectories and Gaussian09[?] for the QM evaluation of energies, gradients and Hessians– to enable QM/MM direct dynamics.

Concluding remarks. It is important to recognize the difference between this chemical dynamics simulation for which *an ensemble of trajectories* are calculated and a molecular dynamics (MD) simulation.[151] For both the atomistic dynamics are obtained by numerically integrating the classical equations of motion. However, for a traditional MD simulation *a single trajectory* is calculated whose motion is assumed to be ergodic. If this trajectory is coupled to a thermal bath,[151] the properties of a constant temperature T canonical ensemble are determined by following the trajectory *vs.* time.

With direct dynamics simulations, we can probe reaction mechanisms

at the atomic level and use the results to interpret the experiments. The simulations also allow a deeper understanding of reaction mechanisms than provided by the widely used reaction path, transition state theory, and Rice-Ramsperger-Kassel-Marcus (RRKM) theory models.[140] By means of direct dynamics simulations it is also possible to discover new reaction pathways and chemical dynamics.

7 Experimental methodology

7.1 Collision Induced Dissociation

The first observations of the products of the collision-induced decomposition (CID) of ions were made in the early days of the development of mass spectrometry and for many years, they were regarde as little more than a nuisance.[152] Over the past 40 years, the CID of ions has been transformed from a nuisance, through being a curiosity to being a major technique used in determining traces of specific components in complex mixtures without the use of chromatography. Modern applications of CID, also referred to as collision activated dissociation (CAD) are detection, identification, and structural analysis of organic molecules, to complex mixture analysis, and to biopolymer sequencing.[153–157] CID has proven extremely useful for the identification and characterization of ions and for complex mixture analysis.[158]

Tandem mass pectrometry (MS/MS) refers to the coupling of two mass spectrometers in time and space with the objective of obtaining further information about the sample in question. Tandem mass spectrometry is based upon the activation of an isolated precursor ion and the analysis of its fragmentation products. In the last decaces, a wide range of strategies and tools have been developed in order to improve the activation of ions and therefore the structural elucidation of molecules.[159,160] The most common ion activation method used in present day instruments remains being CID.[161–163] In a CID experiment, the previously selected and accelerated ions are admitted into the colision cell were they collide with neutral gas targets (typically helium, nitrogen or argon).

When an ion with a high translational energy undergoes an inelastic collision with a neutral, part of the translational energy is converted into internal energy of the ion, leading to subsequent fragmentation.[164] The overall CID process is assumed to occur by a two-step mechanism, where the excitation of the precursors and their fragmentations are separated in time (see 167).

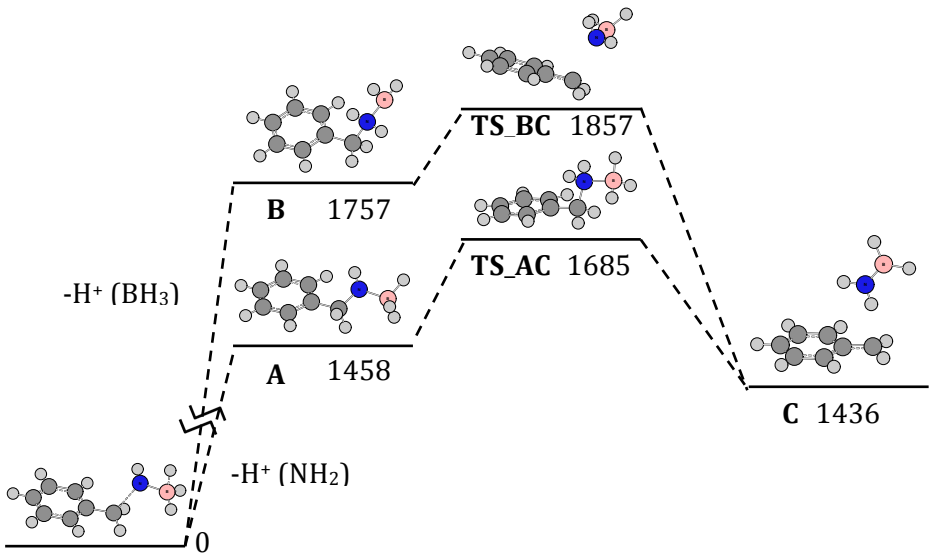


Figure 51: Energy profile of the NH_2 and BH_3deprotonation processes of benzylamine·BH_3. All values in kJ mol^{-1}.

$$m_p^+ + N \rightarrow m_p^{+\ddagger} \rightarrow m_f^+ + m_n \tag{167}$$

where m_p^+, N, $m_p^{+\ddagger}$, m_f^+, and m_n represent the masses for the precursor ion, neutral target, excited precursor ion, and charged and neutral fragments, respectively. In first step a fraction of the ion kinetic energy is transferred into the internal modes of the selected ion. The second part of this mechanism is a unimolecular dissociation of an excited ion, what explains the use of statistical theories such as TST or RRKM to rationalize CID spectra. Fragmentation of the precursor ion can occur if the collision energy is sufficiently high that the ion is excited beyond its threshold for dissociation.

The transfer of kinetic energy to internal energy can be represented by the laws of physics involving a mobile species (ion) and a static target (gas). To simplify the description of such a process, it is more useful to work in the center-of-mass (com9 framework instead of the laboratory reference frame. In the latter, a binary collision is described by the two separate particles involved with their individual position and velocity vectors. The velocities of the ion and neutral gas are stated as velocities relative to each other. The com kinetic energy of the colliding particles, not the laboratory kinetic energy of the ion, is the important parameter determining the nature of the

activation step of the collision process. The totla available energy for the transfer of kinetic energy to iinternal energy is the relative energy (E_{com}) and depends on the collision partners' masses. Equation (168) relates the center of mass and laboratory collision energies:

$$E_{com} = \frac{N}{m_p + N} E_{lab}; \quad E_{lab} = q \cdot V \tag{168}$$

where E_{lab} is the ion's kinetic energy and q and V are the charge and acceleration voltage of the ion, respectively. The CID process is highly depend on the relative masses of the two species. Conservation of energy means that if the relative translational energies of the colliding particles change by a certain amount this energy must appear as internal energy. E_{com} represents the maximum amount of energy that can be converted into internal energy of the precursor ion. This energy, as seen in the equation above, increases with the target's mass, allowing more of the ion's kinetic energy to be converted into internal energy. Furthermore, E_{com} decreases as a function of $1/m_p$, so larger precursor ions have less internal energy available for fragmentation. through the collision process.[165]

All CID processes occurring routinely can be separated into one of two categories based primarily on the translational energy of the precursor ion. For ions of moderate mass (several hundred daltons), *low-energy collisions* occur in the 1–100 eV range of collision energy; and *high-energy collisions* are in the kiloelectronvolt range. Intermediate collision energies (100–1000 eV) do not occur in commonly used tandem mass spectrometers. In the work present in this PhD thesis (Part II) we will be dealing with low-energy collisions.

7.2 Cooks kinetic method

The "kinetic method", intiated 30 years ago by Cooks and co-workers[166–171] is one of the most widely used mass spectrometry technique for the determination of thermochemical quantities in the gas-phase.[23] It provides a relatively easy way to determine a variety of thermochemical values. Enthalpy determinations are relatively straighforward, although entropic effects.

In the "standard kinetic method",[170] the entropy term is assuemd to be zero. The "extended kinetic method"[170] developed by Fenselau,[15,172]

Wedemiotis[16,173,174] and their co-workers employs different instrumental conditions to vary the effective temperature parameter as a thermodynamic temperature to extract apparent entropy differences between pairs of cluster ions. Armentrout proposed an improved statistical treatment of extended kinetic method data to obtain enthalpy differences and the apparent entropies with realistic uncerainties.[17]

To determine the deprotonation thermochemistry of a molecule AH, $AH \rightarrow A^- + H^+$, the kinetic method considers the competitive dissociations of a series of proton bound heterodimers [AHB_i], where B_i is a set of reference acids with known gas-phase acidity values:

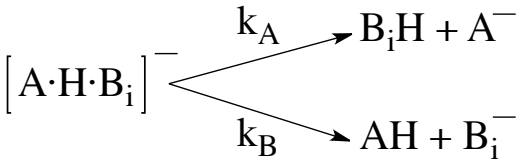


Figure 52: Collision induced dissociations of $[A{\cdot}H{\cdot}B_i]^-$.

[AHB_i] adduct ions can be produced by means of chemical ionization or an electrospray ion source and their dissociations (spontaneous or induced by collision) are analyzed after selection of the adduct ion by the first mass analyzer.

The method generally supposes that the peak intensities $[A]^-$ and $[B_i]^-$ reflect the corresponding rate constants as shown in eq (169). Implicit in this expresion and the followings is the assumption that there is no reverse activation barrier to either dissociation process.

$$\frac{[A]^-}{[B_i]^-} = \frac{k_A}{k_{B_i}} \tag{169}$$

Accoridng to the absolute rate theory, the canonical rate constant k associated with a population of species in thermal equilibrium at a given temperature T is given by eq. (170)

$$k = \left(\frac{k_B T}{h}\right) \exp\left(\frac{-\Delta G}{RT}\right) \tag{170}$$

where k_B is the Boltzmann constant, h the Planck constant, R the gas constant and ΔG is the activation energy of the considered reaction. Then ,

according to eq. (170), the natural logarithm of the peaks ratio may be expressed by eq. (171)

$$\ln\left(\frac{k_A}{k_{B_i}}\right) = \frac{\Delta G}{RT} \tag{171}$$

where ΔG is the free energy of the equilibrium reaction, $A^- + HB_i \leftrightarrow AH + B_i^-$. The free energy of reaction is then related to enthalpy by the standard thermodinamic expresion (172)

$$\Delta G = \Delta(\Delta H) - T\Delta S = \Delta H(AH) - \Delta H(B_iH) - T\Delta S \tag{172}$$

where ΔS is the entropy change for the equilibrium reaction. At this point, it should be emphasized that, since the system cannot be considered at thermal equilibrium, the temperature T is an "effective temperature", T_{eff}. The effective temperature parameter is an empirical parameter, which is closely related to the internal energy of ions actually dissociating in the mass spectrometer.[175,176] However it is not directly related to the mean internal energy of the whole ion population, nor is it necessarily an approximation of the "real" temperature.

What Fenselau and Wesdemiotis realized is that by acquiring $\ln(k_A/K_{B_i})$ data at several values of T_{eff}, they could independently determine $\Delta(\Delta H)$ and Δ. T_{eff} can be varied in several ways: by using collision induced dissociation (CID) and varying the kinetic energy or by changing the collision gas (thereby, changing the laboratory to center of mass energy conversion).

To determine entropic effects, the procedure of Fenselau and Wesdemiotis first combines eq. (171) and eq. (172) to give eq. (173).

$$\ln\left(\frac{k_A}{k_{B_i}}\right) = \left[\frac{\Delta H(AH)}{RT_{eff}} - \frac{\Delta S}{R}\right] - \frac{\Delta H(B_iH)}{RT_{eff}} \tag{173}$$

By plotting $\ln(k_A)/k_{B_i}$ *vs.* the known deprotonation enthalpies, $\Delta H(B_iH)$, one can extract a slope (m_1) of $-1/RT_{eff}$ and an intercept (y_{01}) given by eq. (174).

$$\frac{DG^{app}(AH)}{RT_{eff}} = \frac{\Delta H(AH)}{RT_{eff}} - \frac{\Delta S}{R} \tag{174}$$

where Wesdemiotis introduces the "apparent" free energy of deprotonation (eq. (175))

$$DG^{app}(AH) = \Delta H(AH) - T_{eff}\Delta S \tag{175}$$

The term "apparent" refers to the fact that the ΔS term contains the difference in entropies of dissociation for AH and B_iH, rather than being just the entropy of dissociation of AH. It is useful to note that $DG^{app}(AH)$ is the x intercept of the $\ln(k_A)/k_{B_i}$ *vs.* $\Delta H(B_iH)$ plot.

In a next step, the desired quantities $\Delta H(AH)$ and ΔS are determined by plotting $DG^{app}(AH)/RT_{eff}$ (the intercept y_{01}) *vs.* $1/RT_{eff}$, (the negative of the slope m_1). Eq. (174) shows that the slope of this second plot (m_2) equals $\Delta H(AH)$ and the intercept (y_{02}) equals $-\Delta S/R$.

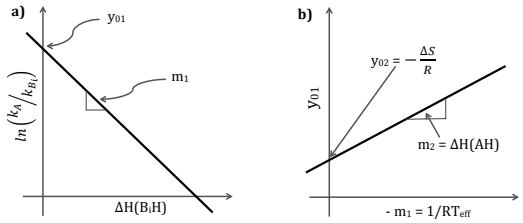


Figure 53: EKM plots for a) $\ln(k_A/k_{B_i})$ *vs.* $\Delta H(B_iH)$ and b) the intercept y_{01} *vs.* the slope m_1 both obtained from plot a).

In 2000, Armentrout[17] pointed out a conceptual error in the previously explained statistical analysis performed to determine $\Delta H(AH)$ and ΔS and proposed "statistically significant ways of handling the same dat to acquire the same information". The problem with the plot in the right side of Fig. 53 is that in the linear regression analysis of this figure, the slope and the intercept used to describe the best fit of the available data are strongly correlated. Therefore, small changes in the slope of the data result in large changes in the y intercept, mirrored in very large uncertainties in the y intercept.

To rigorously remove the correlation between the slope and intercept of a linear regression analysis, one merely needs to plot y *vs.* $x'_i = x_i - x_{avg}$ where x_{avg} is the average deprotonation enthalpy of the reference acids, B_iH. The slope of such a plot is identical to that obtained from plotting $\ln(k_A/K_{B_i})$ *vs.* $\Delta H(B_iH)$, but now the y intercept (y'_{01}) is an interpolated point (and hence much more accurate and precise), rather than being extrapolated. It can be verified that the covariance between y'_0 and m is zero, *i.e.*, the y intercept and slope of this plot are uncorrelated.

For kinetic method data, the slope of this plot (m_1) is again $-1/RT_{eff}$

and the y intercept is given by the expression (176)

$$y'_{01} = \frac{[DG^{app}(AH) - \Delta H_{avg}]}{RT_{eff}} = \frac{\Delta H(AH) - \Delta H_{avg}}{RT_{eff}} - \frac{\Delta S}{R} \qquad (176)$$

Now a plot of the intercept (y'_{01}) *vs.* the negative slope ($-m_1 = 1/RT_{eff}$) will give a slope (m'_2) of $\Delta H(AH) - \Delta H_{avg}$ and a y intercept (y_{02}) of $-\Delta S/R$.

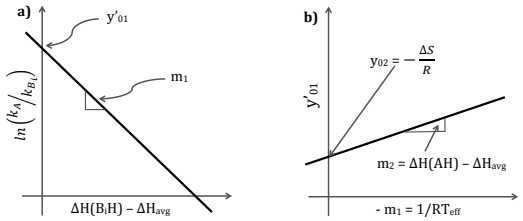


Figure 54: EKM plots for a) $\ln(k_A/k_{B_i})$ *vs.* $\Delta H(B_iH) - \Delta H_{avg}$ and b) the intercept y'_{01} *vs.* the slope m_1 both obtained from plot a).

In spite of much controversy about using the kinetic method to evaluate entropy differences, now it seems widely accpted that the kinetic method can be used to determine enthalpy and entropy differences.[22]

For the experimental acidity values presented here, eighteen compounds with known gas-phase acidites ranging from 1343.5 to 1463.1 kJ mol^{-1} were chosen as the reference acids, B_iH. The ionization technique used to produce the [AHB$_i$] adduct ions was electrospray ionization. The dissociation of the adduct ions after selection by the first mass analyzer was induced by collision with Ar at a nominal pressure of 0.2 mTorr to mantain single collision conditions. The CID experiments were performed using 6 to 13 different center-of-mass collision energies varying from 0.75 to 3.0 eV.

List of Figures

Bibliography

[1] Julia Rehbein and Barry K. Carpenter. Do we fully understand what controls chemical selectivity? *Phys. Chem. Chem. Phys.*, 13:20906–20922, 2011.

[2] Erin Shay, Elide De Gandiaga, and Amy K. Madl. Considerations for the development of health-based surface dust cleanup criteria for beryllium. *Critical Reviews in Toxicology*, 43(3):220–243, 2013.

[3] F. Albert Cotton, Geoffrey Wilkinson, Carlos A. Murillo, and Manfred Bochmann. *Adavanced Inorganic Chemistry*. John Wiley & Sons, New York, sixth edition edition, 1999.

[4] N. N. Greenwood and A. Earnshaw. *Chemistry of the elements*. Pergamon Press, 1984.

[5] Gay-Lussac J. L. *Mem. Phys. Chim. Soc. d'Arcueil*, 2:211, 1809.

[6] Anne Staubitz, Alasdair P. M. Robertson, Matthew E. Sloan, and Ian Manners. Amine and phosphineborane adducts: New interest in old molecules. *Chemical Reviews*, 110(7):4023–4078, 2010.

[7] Timothy J. Clark, Kajin Lee, and Ian Manners. Transition-metal-catalyzed dehydrocoupling: A convenient route to bonds between main-group elements. *Chemistry A European Journal*, 12(34):8634–8648, 2006.

[8] J. Davy. *Phil. Trans*, 30:365, 1812.

[9] Muetterties E. L. *Boron Hydride Chemistry*. Academic Press, New York, 1975.

[10] Gregory C. Welch, Ronan R. San Juan, Jason D. Masuda, and Douglas W. Stephan. Reversible, metal-free hydrogen activation. *Science*, 314(5802):1124–1126, 2006.

[11] Lane C. F. *Aldrichim. Acta*, page 51, 1973.

[12] Anne Staubitz, Alasdair P. M. Robertson, and Ian Manners. Ammonia-borane and related compounds as dihydrogen sources. *Chemical Reviews*, 110(7):4079–4124, 2010.

[13] Kaszynski P., Pakhomov S., Gurskii M. E., Erdyakov S. Y., Starikova Z. A., Lyssenko K. A., Antipin M. Y., Young V. G., and Bubnov Y. N. *J. Org. Chem.*, 74:1709, 2009.

[14] Carboni B. and Monnier L. *Tetrahedron*, 55:1197, 1999.

[15] Xueheng Cheng, Zhuchun Wu, and Catherine Fenselau. Collision energy dependence of proton-bound dimer dissociation: entropy effects, proton affinities, and intramolecular hydrogen-bonding in protonated peptides. *Journal of the American Chemical Society*, 115(11):4844–4848, 1993.

[16] Blas A. Cerda and Chrys Wesdemiotis. Li+, na+, and k+ binding to the dna and rna nucleobases. bond energies and attachment sites from the dissociation of metal ion-bound heterodimers. *Journal of the American Chemical Society*, 118(47):11884–11892, 1996.

[17] P.B. Armentrout. Entropy measurements and the kinetic method: a statistically meaningful approach. *Journal of the American Society for Mass Spectrometry*, 11(5):371 – 379, 2000.

[18] Xubin Zheng and R. Graham Cooks. Thermochemical determinations by the kinetic method with direct entropy correction. *The Journal of Physical Chemistry A*, 106(42):9939–9946, 2002.

[19] Kent M. Ervin. Microcanonical analysis of the kinetic method. the meaning of the apparent entropy. *Journal of the American Society for Mass Spectrometry*, 13(5):435 – 452, 2002.

[20] László Drahos and Károly Vékey. Entropy evaluation using the kinetic method: is it feasible? *Journal of Mass Spectrometry*, 38(10):1025–1042, 2003.

[21] Kent M. Ervin and P. B. Armentrout. Systematic and random errors in ion affinities and activation entropies from the extended kinetic method. *Journal of Mass Spectrometry*, 39(9):1004–1015, 2004.

[22] László Drahos, Csaba Peltz, and Károly Vékey. Accuracy of enthalpy and entropy determination using the kinetic method: are we approaching a consensus? *Journal of Mass Spectrometry*, 39(9):1016–1024, 2004.

[23] Guy Bouchoux. Gas-phase basicities of polyfunctional molecules. part 1: Theory and methods. *Mass Spectrometry Reviews*, 26(6):775–835, 2007.

[24] E. D. Glendening, J. Badenhoop, K. A. E. Reed, J. E. Carpenter, J. A. Bohmann, J. A. Morales, and F. Weinhold. Nbo 5g program, 2001. Theoretical Chemistry Institute, University of Wisconsin, Madison.

[25] K.B. Wiberg. Application of the pople-santry-segal CNDO method to the cyclopropylcarbinyl and cyclobutyl cation and to bicyclobutane. *Tetrahedron*, 24(3):1083 – 1096, 1968.

[26] R. F. W. Bader. *Atoms in Molecules. A Quantum Theory*. Clarendon Press, Oxford, 1990.

[27] A. D. Becke and K. E. Edgecombe. A simple measure of electron localization in atomic and molecular systems. *The Journal of Chemical Physics*, 92(9):5397–5403, 1990.

[28] B. Silve and A. Savin. Calssification of chemical bonds based on topological analysis of electron localization functions. *Nature*, 371(6499):683–686, 1994.

[29] Todd A. Keith. Aimall (version 11.12.19), 2011. TK Gristmill Software, Overland Park KS, USA.

[30] S. Noury, X Krokidis, F. Fuster, and B. Silvi. Topmod package, 1997. Universite Pierre et Marie Curie.

[31] Cameron Jones, George A. Koutsantonis, and Colin L. Raston. Lewis base adducts of alane and gallane. *Polyhedron*, 12(15):1829 – 1848, 1993.

[32] Colin L. Raston. Recent developments in the chemistry of alane (alh3) and gallane (gah3). *Journal of Organometallic Chemistry*, 475(12):15 – 24, 1994.

[33] Michael G. Gardiner and Colin L. Raston. Advances in the chemistry of lewis base adducts of alane and gallane. *Coordination Chemistry Reviews*, 166(0):1 – 34, 1997.

[34] Danan Dou, Douglas R. Ketchum, Ewan J. M. Hamilton, Pierre A. Florian, Karl E. Vermillion, Philip J. Grandinetti, and Sheldon G. Shore. Reactions of aluminum hydride derivatives with ammoniaborane: a new approach toward aln/bn materials. *Chemistry of Materials*, 8(12):2839–2842, 1996.

[35] Philip C. Andrews, Michael G. Gardiner, Colin L. Raston, and Vicki-Anne Tolhurst. Structural aspects of tertiary amine adducts of alane and gallane. *Inorganica Chimica Acta*, 259(12):249 – 255, 1997.

[36] Fiona M. Elms, Michael G. Gardiner, George A. Koutsantonis, Colin L. Raston, Jerry L. Atwood, and Kerry D. Robinson. Tertiary phosphine adducts of alane and gallane. *Journal of Organometallic Chemistry*, 449(12):45 – 52, 1993.

[37] Nguyen Vinh-Son, Saartje Swinnen, Myrna H. Matus, Minh Tho Nguyen, and David A. Dixon. The effect of the nh2 substituent on nh3: hydrazine as an alternative for ammonia in hydrogen release in the presence of boranes and alanes. *Phys. Chem. Chem. Phys.*, 11:6339–6344, 2009.

[38] Daniel J. Grant and David A. Dixon. Thermodynamic properties of molecular borane phosphines, alane amines, and phosphine alanes and the [bh4-][ph4+], [alh4-][nh4+], and [alh4-][ph4+] salts for chemical hydrogen storage systems from ab initio electronic structure theory. *The Journal of Physical Chemistry A*, 109(44):10138–10147, 2005.

[39] David Pugh, Peter Marchand, Ivan P. Parkin, and Claire J. Carmalt. Group 13 -ketoiminate compounds: Gallium hydride derivatives as molecular precursors to thin films of ga2o3. *Inorganic Chemistry*, 51(11):6385–6395, 2012.

[40] Everett M. Marlett and Won Suh Park. Dimethylethylamine alane and n-methylpyrrolidine alane. a convenient synthesis of alane, a useful selective reducing agent in organic synthesis. *The Journal of Organic Chemistry*, 55(9):2968–2969, 1990.

[41] Colin L. Raston, Anna F.H. Siu, Carolyn J. Tranter, and David J. Young. Functional group reductions with lewis base adducts of gallane. *Tetrahedron Letters*, 35(32):5915 – 5918, 1994.

[42] George A. Koutsantonis, Fu Chin Iee, and Colin L. Raston. Hydride-bridged heterobimetallic complexes of gallium and zinc: the first x-ray

structural determination of the gah4-moiety. *J. Chem. Soc., Chem. Commun.*, pages 1975–1976, 1994.

[43] Roland A. Fischer, Joachim Behm, Thomas Priermeier, and Wolfgang Scherer. Transition-metal-substituted volatile alanes, gallanes, and indanes: Synthesis and structure. *Angewandte Chemie International Edition in English*, 32(5):746–748, 1993.

[44] Karl Khan, Colin L. Raston, John E. McGrady, Brian W. Skelton, and Allan H. White. Hydride-bridged heterobimetallic complexes of zirconium and aluminum. *Organometallics*, 16(15):3252–3254, 1997.

[45] Ghenwa Bouhadir, Abderrahmane Amgoune, and Didier Bourissou. Chapter 1 - phosphine-boranes and related ambiphilic compounds: Synthesis, structure, and coordination to transition metals. volume 58 of *Advances in Organometallic Chemistry*, pages 1 – 107. Academic Press, 2010.

[46] P.J. Linstrom and W. G. Mallard, editors. *NIST Chemistry Webbook. Standard Reference Database Number 69*. National Institute of Standards and Technology, Gaithersburg MD, 20899, release june 2005 edition, 2012.

[47] Kent M. Ervin and W. Carl Lineberger. Photoelectron spectroscopy of phosphorus hydride anions. *The Journal of Chemical Physics*, 122(19):–, 2005.

[48] Marcela Hurtado, Manuel Ynez, Rebeca Herrero, Andrs Guerrero, JuanZ. Dvalos, Jos-LuisM. Abboud, Brahim Khater, and Jean-Claude Guillemin. The ever-surprising chemistry of boron: Enhanced acidity of phosphineboranes. *Chemistry A European Journal*, 15(18):4622–4629, 2009.

[49] Christoph Loschen, Katrin Voigt, Jan Frunzke, Axel Diefenbach, Michael Diedenhofen, and Gernot Frenking. Theoretical studies of inorganic compounds. 19 1) quantum chemical investigations of the phosphane complexes x3b-py3 and x3al-py3 (x = h, f, cl; y = f, cl, me, cn). *Zeitschrift fr anorganische und allgemeine Chemie*, 628(6):1294–1304, 2002.

[50] Margaret Czerw, Alan S. Goldman, and Karsten Krogh-Jespersen. Addition of ammonia to alh3 and bh3. why does only aluminum form 2:1 adducts? *Inorganic Chemistry*, 39(2):363–369, 2000.

[51] Manuel Yez, Otilia M, Ibon Alkorta, and Jos Elguero. Can conventional bases and unsaturated hydrocarbons be converted into gas-phase superacids that are stronger than most of the known oxyacids? the role of beryllium bonds. *Chemistry A European Journal*, 19(35):11637–11643, 2013.

[52] T. Mark McCleskey and Brian L. Scott. Beryllium and strong hydrogen bonds. *Journal of Occupational and Environmental Hygiene*, 6(12):751–757, 2009. PMID: 19894176.

[53] Beryllium. In J. P. Hammel and G. G. Hawley, editors, *The Encyclopedia of Chemistry*. Renihold, New York, 1973.

[54] Tammy P. Taylor, Mei Ding, Deborah S. Ehler, Trudi M. Foreman, John P. Kaszuba, and Nancy N. Sauer. Beryllium in the environment: A review. *Journal of Environmental Science and Health, Part A*, 38(2):439–469, 2003.

[55] Otilia M, Manuel Yáñez, Michle Decouzon, Jean-Franois Gal, Pierre-Charles Maria, and Jean-Claude Guillemin. Gas-phase basicity and acidity trends in ,-unsaturated amines, phosphines, and arsines. *Journal of the American Chemical Society*, 121(19):4653–4663, 1999.

[56] Jean-Claude Guillemin, Michle Decouzon, Pierre-Charles Maria, Jean-Franois Gal, Otilia M, and Manuel Yez. Gas-phase basicities and acidities of ethyl-, vinyl-, and ethynylarsine. an experimental and theoretical study. *The Journal of Physical Chemistry A*, 101(49):9525–9530, 1997.

[57] Jean-Claude Guillemin, El Hassan Riague, Jean-Franois Gal, Pierre-Charles Maria, Otilia M, and Manuel Yez. Acidity trends in ,-unsaturated sulfur, selenium, and tellurium derivatives: Comparison with c-, si-, ge-, sn-, n-, p-, as-, and sb-containing analogues. *Chemistry A European Journal*, 11(7):2145–2153, 2005.

[58] Lide D. R. *CRC Handbook of Chemistry and Physics*. CRC Press, Boca Raton, FL, 69th edition, 2008.

[59] Plumley J. A. and Evanseck J. D. *J. Phys. Chem. A*, 113:5985, 2009.

[60] Brinck T., Murray J. S., and Politzer P. *Inorg. Chem.*, 32:2622, 1993.

[61] Bessac F. and Frenking G. *Inorg. Chem.*, 42:7990, 2003.

[62] Baban J. A., Marti V. P. J., and Roberts B. P. *J. Chem. Soc., Perkin Trans. 2*, page 1723, 1985.

[63] Baban J. A. and Roberts B. P. *J. Chem. Soc., Chem. Commun.*, page 1224, 1983.

[64] Green I. G. and Roberts B. P. *J. Chem. Soc., Perkin Trans. 2*, page 1597, 1986.

[65] Baban J. A., Marti V. P. J., and Roberts B. P. *J. Chem. Res.*, page 90, 1985.

[66] Mok P. L. H., Roberts B. P., and McKetty P. T. *J. Chem. Soc., Perkin Trans. 2*, page 665, 1993.

[67] Baban J. A. and Roberts B. P. *J. Chem. Soc., Perkin Trans. 2*, page 1195, 1988.

[68] Paul V. and Roberts B. P. *J. Chem. Soc., Perkin Trans. 2*, page 1895, 1988.

[69] Barton D. H. R. and Jacob M. *Tetrahedron Lett.*, 39:1331, 1998.

[70] Little R. D. and Nishiguchi G. A. *Stud. Nat. Prod. Chem.*, 35:3, 2008.

[71] Nishiguchi G. A. and Little R. D. *J. Org. Chem.*, 70:5249, 2005.

[72] Florent J.-C. and Monneret C. *Glycoscience*, page 231, 2001.

[73] Studer A. and Amrein S. *Synthesis*, page 835, 2002.

[74] Gilbert B. C. and Parsons A. F. *J. Chem. Soc., Perkin Trans. 2*, page 367, 2002.

[75] Detlef Schöder and Helmut Schwarz. Generation, stability, and reactivity of small, multiply charged ions in the gas phase. *The Journal of Physical Chemistry A*, 103(37):7385–7394, 1999. and references therein.

[76] Riccardo Spezia, Alvaro Cimas, Marie-Pierre Gaigeot, Jean-Yves Salpin, Kihyung Song, and William L. Hase. Collision induced dissociation of doubly-charged ions: Coulomb explosion vs. neutral loss in [ca(urea)]2+ gas phase unimolecular reactivity via chemical dynamics simulations. *Phys. Chem. Chem. Phys.*, 14:11724–11736, 2012.

[77] Samy O. Meroueh, Yanfei Wang, and William L. Hase. Direct dynamics simulations of collision- and surface-induced dissociation of n-protonated glycine. shattering fragmentation. *The Journal of Physical Chemistry A*, 106(42):9983–9992, 2002.

[78] Oussama Meroueh and William L. Hase. Collisional activation of small peptides. *The Journal of Physical Chemistry A*, 103(20):3981–3990, 1999.

[79] Yannick Jeanvoine, Marie-Pierre Gaigeot, William L. Hase, Kihyung Song, and Riccardo Spezia. Collision induced dissociation of protonated urea with n2: Effects of rotational energy on reactivity and energy transfer via chemical dynamics simulations. *International Journal of Mass Spectrometry*, 308(23):289 – 298, 2011. Eyer Special Issue.

[80] Xiche Hu, William L. Hase, and Tony Pirraglia. Vectorization of the general monte carlo classical trajectory program venus. *Journal of Computational Chemistry*, 12(8):1014–1024, 1991.

[81] Hase W. L., Duchovic R. J., Hu X., Komornicki A., Lim K. F., Lu D.-H., Peslherbe G. H., Swamy K. N., Linde S. R. V., Varandas A., Wang H., and Wolf R. J. *QCPE*, 16:671, 1996.

[82] Ane Eizaguirre, Otilia Mo, Manuel Yanez, and Jean-Yves Salpin. Modeling the interactions between peptide functions and sr2+: formamide-sr2+ reactions in the gas phase. *Phys. Chem. Chem. Phys.*, 13:18409–18417, 2011.

[83] Ane Eizaguirre, Otilia Mo, Manuel Yanez, Jean-Yves Salpin, and Jeanine Tortajada. Modelling peptide-metal dication interactions: formamide-ca2+ reactions in the gas phase. *Org. Biomol. Chem.*, 10:7552–7561, 2012.

[84] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, T. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, O. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski,

and D. J. Fox. Gaussian 09 Revision C.01. Gaussian Inc. Wallingford CT 2010.

[85] William C. Swope, Hans C. Andersen, Peter H. Berens, and Kent R. Wilson. A computer simulation method for the calculation of equilibrium constants for the formation of physical clusters of molecules: Application to small water clusters. *The Journal of Chemical Physics*, 76(1), 1982.

[86] T. Raz and R. D. Levine. On the shattering of clusters by surface impact heating. *The Journal of Chemical Physics*, 105(18), 1996.

[87] David G. Schultz and Luke Hanley. Shattering of sime3+ during surface-induced dissociation. *The Journal of Chemical Physics*, 109(24), 1998.

[88] John A. Burroughs, Samuel B. Wainhaus, and Luke Hanley. Impulsive excitation of fecp+2 and sime+3 during surfaceinduced dissociation at organic multilayers. *The Journal of Chemical Physics*, 103(15), 1995.

[89] Oussama Meroueh and William L. Hase. Effect of surface stiffness on the efficiency of surface-induced dissociation. *Phys. Chem. Chem. Phys.*, 3:2306–2314, 2001.

[90] Kihyung Song, Oussama Meroueh, and William L. Hase. Dynamics of cr(co)6+ collisions with hydrogenated surfaces. *The Journal of Chemical Physics*, 118(6), 2003.

[91] Riccardo Spezia, Jean-Yves Salpin, Marie-Pierre Gaigeot, William L. Hase, and Kihyung Song. Protonated urea collision-induced dissociation. comparison of experiments and chemical dynamics simulations. *The Journal of Physical Chemistry A*, 113(50):13853–13862, 2009. PMID: 19886650.

[92] John A. Pople, Martin HeadGordon, Douglas J. Fox, Krishnan Raghavachari, and Larry A. Curtiss. Gaussian1 theory: A general procedure for prediction of molecular energies. *The Journal of Chemical Physics*, 90(10), 1989.

[93] Larry A. Curtiss, Christopher Jones, Gary W. Trucks, Krishnan Raghavachari, and John A. Pople. Gaussian1 theory of molecular energies for secondrow compounds. *The Journal of Chemical Physics*, 93(4), 1990.

[94] Larry A. Curtiss, Krishnan Raghavachari, Gary W. Trucks, and John A. Pople. Gaussian2 theory for molecular energies of first and secondrow compounds. *The Journal of Chemical Physics*, 94(11), 1991.

[95] Larry A. Curtiss, Krishnan Raghavachari, Paul C. Redfern, Vitaly Rassolov, and John A. Pople. Gaussian-3 (g3) theory for molecules containing first and second-row atoms. *The Journal of Chemical Physics*, 109(18), 1998.

[96] Larry A. Curtiss, Paul C. Redfern, and Krishnan Raghavachari. Gaussian-4 theory. *The Journal of Chemical Physics*, 126(8):–, 2007.

[97] Larry A. Curtiss, Paul C. Redfern, Krishnan Raghavachari, and John A. Pople. Gaussian-3x (g3x) theory: Use of improved geometries, zero-point energies, and hartreefock basis sets. *The Journal of Chemical Physics*, 114(1):108–117, 2001.

[98] Asger Halkier, Trygve Helgaker, Poul Jrgensen, Wim Klopper, and Jeppe Olsen. Basis-set convergence of the energy in molecular hartreefock calculations. *Chemical Physics Letters*, 302(56):437 – 446, 1999.

[99] David Feller. The use of systematic sequences of wave functions for estimating the complete basis set, full configuration interaction limit in water. *The Journal of Chemical Physics*, 98(9):7059–7071, 1993.

[100] Thom H. Dunning. Gaussian basis sets for use in correlated molecular calculations. i. the atoms boron through neon and hydrogen. *The Journal of Chemical Physics*, 90(2):1007–1023, 1989.

[101] David E. Woon and Thom H. Dunning. Gaussian basis sets for use in correlated molecular calculations. iii. the atoms aluminum through argon. *The Journal of Chemical Physics*, 98(2):1358–1371, 1993.

[102] Angela K. Wilson, David E. Woon, Kirk A. Peterson, and Thom H. Dunning. Gaussian basis sets for use in correlated molecular calculations. ix. the atoms gallium through krypton. *The Journal of Chemical Physics*, 110(16):7667–7676, 1999.

[103] Larry A. Curtiss, Paul C. Redfern, and Krishnan Raghavachari. Assessment of gaussian-3 and density-functional theories on the g3/05 test set of experimental energies. *The Journal of Chemical Physics*, 123(12):–, 2005.

[104] Daudel R. *Can. J. Chem.*, 52:1310, 1974.

[105] Bader R. F. W. *J. Am. Chem. Soc.*, 97:7391, 1975.

[106] Daudel R. *J. Chem. Phys.*, 23:2080, 1955.

[107] Ponec R. *J. Math. Chem.*, 21:323, 1997.

[108] Ponec R. *J. Comput. Chem.*, 20:760, 1999.

[109] Gillespie R. J. *Gillespie, R. J.; Robinson, E. A. Angew. Chem., Int. Ed. Engl.*, 35:495, 1996.

[110] Gillespie R. J. and Popelier P. L. A. *Chemical Bonding and Molecular Geometries: From Lewis to Electron Densities*. Oxford University Press, New York, 2001.

[111] Bernard Silvi. The synaptic order: a key concept to understand multicenter bonding. *Journal of Molecular Structure*, 614(13):3 – 10, 2002.

[112] McWeeney R. *Methods of Molecular Quantum Mechanics*. Academic, New York, 2nd edition, 1976.

[113] Diner S. and Claverie P. *Localization and Delocalization in Quantum Chemistry*. Reidel, Dordrecht, The Netherlands, 1976. Vol. II.

[114] Bernard Silvi. How topological partitions of the electron distributions reveal delocalization. *Phys. Chem. Chem. Phys.*, 6:256–260, 2004.

[115] Xavier Fradera, Maggie A. Austen, and Richard F. W. Bader. The lewis model and beyond. *The Journal of Physical Chemistry A*, 103(2):304–314, 1999.

[116] Iupac.compendium of chemical terminology, 1997. Compiled by A. D. McNaught and A. Wilkinson.

[117] Oscar Knefler Rice and Herman C. Ramsperger. Theories of unimolecular gas reactions at low pressures. *Journal of the American Chemical Society*, 49(7):1617–1629, 1927.

[118] L. S. Kassel. Studies in homogeneous gas reactions. i. *The Journal of Physical Chemistry*, 32(2):225–242, 1927.

[119] R. A. Marcus and O. K. Rice. *Journal of physical and colloid chemistry*, 55:894, 1951.

[120] C. N. Hinshelwood. On the theory of unimolecular reactions. *Proceedings of the Royal Society of London. Series A*, 113(763):230–233, 1926.

[121] E. Wigner. Calculation of the rate of elementary association reactions. *The Journal of Chemical Physics*, 5(9):720–725, 1937.

[122] J. O. Hirschfelder and E. Wigner. Some quantummechanical considerations in the theory of reactions involving an activation energy. *The Journal of Chemical Physics*, 7(8):616–628, 1939.

[123] H. M. Rosenstock, M. B. Wallenstein, A. L. Wahrhaftig, and Henry Eyring. Absolute rate theory for isolated systems and the mass spectra of polyatomic molecules. *Proceedings of the National Academy of Sciences*, 38(8):667–678, 1952.

[124] Tomas Baer and William L. Hase. *Unimolecular Reaction Dynamics. Theory and experiments*. Oxford University Press, New York, 1986.

[125] Bruce C. Garrett and Donald G. Truhlar. Criterion of minimum state density in the transition state theory of bimolecular reactions. *The Journal of Chemical Physics*, 70(4):1593–1598, 1979.

[126] Xiche Hu and William L. Hase. Modification of the duchovichaseschlegel potential energy function for h+ch3ch4. comparison of canonical variational transition state theory, trajectory, and experimental association rate constants. *The Journal of Chemical Physics*, 95(11):8073–8082, 1991.

[127] Ling Zhu and William L. Hase. Comparison of models for calculating the RRKM unimolecular rate constant k(e, j). *Chemical Physics Letters*, 175(12):117 – 124, 1990.

[128] C. H. Townes and A. L. Schadow. *Microwave Spectrometry*. McGraw-Hill, New York, 1955.

[129] D. C. Tardy, B. S. Rabinovitch, and G. Z. Whitten. Vibrationrotation energylevel density calculations. *The Journal of Chemical Physics*, 48(3):1427–1429, 1968.

[130] G. Z. Whitten and B. S. Rabinovitch. Approximation for rotationvibration energy level sums. *The Journal of Chemical Physics*, 41(6):1883–1883, 1964.

[131] G. Z. Whitten and B. S. Rabinovitch. Accurate and facile approximation for vibrational energylevel sums. *The Journal of Chemical Physics*, 38(10):2466–2473, 1963.

[132] Grigoriy Vayner, Srirangam V. Addepalli, Kihyung Song, and William L. Hase. Post-transition state dynamics for propene ozonolysis: Intramolecular and unimolecular dynamics of molozonide. *The Journal of Chemical Physics*, 125(1):–, 2006.

[133] Upakarasamy Lourderaj, Kyoyeon Park, and William L. Hase. Classical trajectory simulations of post-transition state dynamics. *International Reviews in Physical Chemistry*, 27(3):361–403, 2008.

[134] L. Zhu and W. L. Hase. *Quantum Chemistry Program Exchange*, 14:664.

[135] Don L. Bunker. Monte carlo calculation of triatomic dissociation rates. i. n2o and o3. *The Journal of Chemical Physics*, 37(2), 1962.

[136] Normand C. Blais and Don L. Bunker. Monte carlo calculations. ii. the reactions of alkali atoms with methyl iodide. *The Journal of Chemical Physics*, 37(11), 1962.

[137] DON L. BUNKER. Classical trajectory methods. In SIDNEY FERNBACH BERNI ALDER and MANUEL ROTENBERG, editors, *Atomic and Molecular Scattering*, volume 10 of *Methods in Computational Physics: Advances in Research and Applications*, pages 287 – 325. Elsevier, 1971.

[138] Goldstein H. *Classical Mechanics*. Addison-Wesley, Menlo Park, CA, 1950.

[139] U. Burkert and N. L. Allinger. *Molecular Mechanics*. Am. Chemical Soc., 1982.

[140] J. I. Steinfeld, J. S. Francisco, and W. L. Hase. *Chemical Kinetics and Dynamics*. Prentice-Hall, 2nd edition, 1999.

[141] Angels Gonzalez-Lafont, Thanh N. Truong, and Donald G. Truhlar. Direct dynamics calculations with nddo (neglect of diatomic differential overlap) molecular orbital theory with specific reaction parameters. *The Journal of Physical Chemistry*, 95(12):4618–4627, 1991.

[142] Bolton K., Hase W. L., and Peslherbe G. H. Direct dynamics simulations of reactive systems. In Thomson D. L., editor, *Multidimensional Molecular Dynamics Methods.*, pages 143–189. World Scientific, London, 1998.

[143] Sun L. and Hase W. L. Born-oppenheimer direct dynamics classical trajectory simulations. *Rev. Comput. Chem.*, 19:79–146, 2003.

[144] Loup Verlet. Computer "experiments" on classical fluids. i. thermodynamical properties of lennard-jones molecules. *Phys. Rev.*, 159:98–103, Jul 1967.

[145] Don L. Bunker and Elizabeth A. Goring-Simpson. Alkali-methyl iodide reactions. *Faraday Discuss. Chem. Soc.*, 55:93–99, 1973.

[146] Lipeng Sun and William L. Hase. Ab initio direct dynamics trajectory simulation of c2h5fc2h4+hf product energy partitioning. *The Journal of Chemical Physics*, 121(18), 2004.

[147] Charles Doubleday, Kim Bolton, Gilles H. Peslherbe, and William L. Hase. Direct dynamics simulation of the lifetime of trimethylene. *Journal of the American Chemical Society*, 118(41):9922–9931, 1996.

[148] Lipeng Sun and William L. Hase. Comparisons of classical and wigner sampling of transition state energy levels for quasiclassical trajectory chemical dynamics simulations. *The Journal of Chemical Physics*, 133(4):–, 2010.

[149] M. Kaupp, P. v. R. Schleyer, H. Stoll, and H. Preuss. Pseudopotential approaches to ca, sr, and ba hydrides. why are some alkaline earth mx2 compounds bent? *The Journal of Chemical Physics*, 94(2):1360–1366, 1991.

[150] S.F. Boys and F. Bernardi. The calculation of small molecular interactions by the differences of separate total energies. some procedures with reduced errors. *Molecular Physics*, 19(4):553–566, 1970.

[151] Allen M. P. and Tildesley D. J. *Computer simulation of liquids.* Clarendon, Oxford, UK, 1987.

[152] R. Graham Cooks. Special feature: Historical. collision-induced dissociation: Readings and commentary. *Journal of Mass Spectrometry*, 30(9):1215–1221, 1995.

[153] J. H. Beynon, R. G. Cooks, J. W. Amy, W. E. Baitinger, and T. Y. Ridley. Design and performance of a mass-analyzed ion kinetic energy (mike) spectrometer. *Analytical Chemistry*, 45(12):1023A–1031A, 1973.

[154] J. H. Futrell and Tiernan T. O. Tandem mass spectrometric studies of ino-molecule reactions. In J. L. Franklin, editor, *Ion-Molecule Reactions*, volume 2, pages 485–549. Plenum Press, New York, 1972.

[155] K.R. Jennings. Collision-induced decompositions of aromatic molecular ions. *International Journal of Mass Spectrometry and Ion Physics*, 1(3):227 – 235, 1968.

[156] F. W. McLafferty, P. F. Bente, Richard. Kornfeld, Shih-Chuan. Tsai, and Ian. Howe. Metastable ion characteristics. xxii. collisional activation spectra of organic ions. *Journal of the American Chemical Society*, 95(7):2120–2129, 1973.

[157] Timothy Wachs, Paul F. Bente III, and F.W. McLafferty. Simple modification of a commercial mass spectrometer for metastable data collection. *International Journal of Mass Spectrometry and Ion Physics*, 9(3):333 – 341, 1972.

[158] J. Mitchell Wells and Scott A. McLuckey. *CollisionInduced Dissociation (CID) of Peptides and Proteins*, volume 402. 2005.

[159] Lekha Sleno and Dietrich A. Volmer. Ion activation methods for tandem mass spectrometry. *Journal of Mass Spectrometry*, 39(10):1091–1112, 2004.

[160] Marija Mentinova and ScottA. McLuckey. Intra- and inter-molecular cross-linking of peptide ions in the gas phase: Reagents and conditions. *Journal of The American Society for Mass Spectrometry*, 22(5):912–921, 2011.

[161] Scott A. McLuckey. Principles of collisional activation in analytical mass spectrometry. *Journal of the American Society for Mass Spectrometry*, 3(6):599 – 614, 1992.

[162] Anil K. Shukla and Jean H. Futrell. Tandem mass spectrometry: dissociation of ions by collisional activation. *Journal of Mass Spectrometry*, 35(9):1069–1090, 2000.

[163] Keith R Jennings. The changing impact of the collision-induced decomposition of ions on mass spectrometry. *International Journal of Mass Spectrometry*, 200(13):479 – 493, 2000. Volume 200: The state of the field as we move into a new millenium.

[164] Levsen K. In *Fundamental Aspects of Organic Mass Spectrometry.* Verlag Chimie, Weinheim, 1978. 138 (collision process), 92 (energy dependence of CID products).

[165] Busch K. L., Glish G. L., and McLuckey S. A. Techniques and applications of tandem mass spectrometry. In *Mass Spectrometry/Mass Spectrometry.* VCH, New York, 1988. 64 (Ecom), 84 (pressure effects).

[166] R. G. Cooks and T. L. Kruger. Intrinsic basicity determination using metastable ions. *Journal of the American Chemical Society*, 99(4):1279–1281, 1977.

[167] S. A. McLuckey, D. Cameron, and R. G. Cooks. Proton affinities from dissociations of proton-bound dimers. *Journal of the American Chemical Society*, 103(6):1313–1317, 1981.

[168] R. Graham Cooks, Jeffrey S. Patrick, Tapio Kotiaho, and Scott A. McLuckey. Thermochemical determinations by the kinetic method. *Mass Spectrometry Reviews*, 13(4):287–339, 1994.

[169] R. Graham Cooks and Philip S. H. Wong. Kinetic method of making thermochemical determinations: advances and applications. *Accounts of Chemical Research*, 31(7):379–386, 1998.

[170] R. G. Cooks, J. T. Koskinen, and P. D. Thomas. The kinetic method of making thermochemical determinations. *Journal of Mass Spectrometry*, 34(2):85–92, 1999.

[171] Zheng X., Cooks R. G., Augusti R., and Tao W. A. The kinetic method: Thermochemical determinations and chiral analysis. In Gross M. L., Caprioli R. M:, and Armentrout P., editors, *Encyclopedia of mass spectrometry, Vol. 1*, pages 350–362. Elsevier, Amsterdam, 2003.

[172] Zhuchun Wu, Catherine Fenselau, and R. Graham Cooks. Gas-phase basicities and proton affinities of lysine and histidine measured from the dissociation of proton-bound dimers. *Rapid Communications in Mass Spectrometry*, 8(9):777–780, 1994.

[173] Blas A. Cerda, Sophie Hoyau, Gilles Ohanessian, and Chrys Wesdemiotis. Na+ binding to cyclic and linear dipeptides. bond energies, entropies of na+ complexation, and attachment sites from the dissociation of na+-bound heterodimers and ab initio calculations. *Journal of the American Chemical Society*, 120(10):2437–2448, 1998.

[174] Blas A. Cerda and Chrys Wesdemiotis. Gas phase copper(i) ion affinities of valine, lysine, and arginine based on the dissociation of cu+-bound heterodimers at varying internal energies. *International Journal of Mass Spectrometry*, 185187:107 – 116, 1999.

[175] Lszl Drahos and Kroly Vkey. How closely related are the effective and the real temperature. *Journal of Mass Spectrometry*, 34(2):79–84, 1999.

[176] Kent M Ervin. Microcanonical analysis of the kinetic method.: The meaning of the effective temperature. *International Journal of Mass Spectrometry*, 195196(0):271 – 284, 2000.

Appendix 1

Table 12: Stabilization free energy of neutral ($\Delta_r G_1^0$) and deprotonated ($\Delta_r G_2^0$) amines upon BH_3 association. G4-based values. All the values in kJ mol^{-1}.

Amine	$\Delta_r G_1^0$	$\Delta_r G_2^0$
ammonia	-77.9	-273.1
methylamine	-96.3	-290.1
dimethylamine	-107.8	-276.0
allylamineine	-95.0	-209.1
cyclopropylamine	-71.2	-251.3
benzylamine	-112.4	-263.2
aziridine	-97.7	-265.5
propargylamine	-90.2	-267.9
trifluoroethylamine	-78.5	-257.3
aniline	-57.7	-209.1

Table 13: Characteristics of the B–N bond in amine-boranes in terms of atomic hybrids, Wiberg bond index, and electron density (ρ_b(B–N), a.u.) at the corresponding BCP.

	Bond		Wiberg bond index		$\rho_b(N-B)$	
Amine	neutral	anion	neutral	anion	neutral	anion
Ammonia	82% **N**(35%s + 65% p)+ 18% **B**(16%s + 84% p)	75% **N**(40%s + 60%p) + 25% **B**(23%s + 77%p)	0.618	0.816	0.105	0.147
Methylamine	82% **N**(34%s + 66% p)+ 18% **B**(16%s + 84% p)	75% **N**(39%s + 61%p) + 25% **B**(23%s + 77%p)	0.612	0.785	0.112	0.151
Dimethylamine	82% **N**(32%s + 68% p)+ 18% **B**(16%s + 84% p)	76% **N**(37%s + 63%p) + 24% **B**(22%s + 78%p)	0.592	0.745	0.114	0.150
Allylamine	82% **N**(33%s + 67% p)+ 18% **B**(16%s + 84% p)	76% **N**(38%s + 62%p) + 24% **B**(22%s + 78%p)	0.610	0.769	0.112	0.148
Cyclopropylamine	82% **N**(33%s + 67% p)+ 18% **B**(16%s + 84% p)	76% **N**(40%s + 60%p) + 24% **B**(22%s + 78%p)	0.598	0.779	0.102	0.147
Benzylamine	LP(**N**)→ 2p (**B**)[a] 1079	76% **N**(39%s + 61%p) + 24% **B**(22%s + 78%p)	0.613	0.762	0.113	0.147
Aziridine	82% **N**(40%s + 60% p)+ 18% **B**(17%s + 83% p)	76% **N**(39%s + 61%p) + 24% **B**(22%s + 78%p)	0.613	0.760	0.116	0.147
Trifluoroethylamine	82% **N**(33%s + 67% p)+ 18% **B**(16%s + 84% p)	76% **N**(38%s + 62%p) + 24% **B**(22%s + 78%p)	0.593	0.761	0.108	0.146
Aniline	LP(**N**)→ 2p (**B**)[a] 1079	77% **N**(43%s + 57%p) + 23% **B**(21%s + 79%p)	0.608	0.743	0.109	0.140

[a]The value reported corresponds to the second order interaction energy, in kJ mol^{-1}, between the N lone pair (LP(**N**)) and the empty $2p$ orbital of **B** atom.

Table 14: Characteristics of the N–Al bond in amine-alanes in terms of atomic hybrids, Wiberg bond index, and electron density (ρ_b(N–Al), a.u.) at the corresponding BCP.

	Bond		Wiberg bond index		$\rho_b(N-Al)$	
R	neutral	anion	neutral	anion	neutral	anion
H	92% **N**(22%s + 78% p)+ 8% **Al**(14%s + 84% p)	87% **N**(42%s + 58%p) + 13% **Al**(25%s + 74%p)	0.290	0.510	0.044	0.072
CH_3	92% **N**(30%s + 70% p)+ 8% **Al**(15%s + 84% p)	88% **N**(39%s + 61%p) + 12% **Al**(25%s + 74%p)	0.279	0.484	0.047	0.073
c-C_3H_5	92% **N**(29%s + 71% p)+ 8% **Al**(15%s + 84% p)	88% **N**(38%s + 62%p) + 12% **Al**(24%s + 74%p)	0.276	0.465	0.046	0.071
C_6H_5	LP(**N**)→3p(**Al**)[a] 247.2	LP(**N**)→ 3p (**Al**)[a] 499.3	0.255	0.433	0.043	0.067

[a]The value reported corresponds to the second order interaction energy, in kJ mol^{-1}, between the N lone pair (LP(**N**)) and the empty 3p orbital of **Al** atom.

Table 15: Characteristics of the P–Al bond in phosphine-alanes in terms of atomic hybrids, Wiberg bond index, and electron density (ρ_b(P–Al), a.u.) at the corresponding BCP.

	Bond		Wiberg bond index		$\rho_b(P-Al)$	
R	neutral	anion	neutral	anion	neutral	anion
H	87% **P**(39%s + 61% p)+ 13% **Al**(12%s + 87% p)	76% **P**(23%s + 77%p) + 24% **Al**(22%s + 77%p)	0.457	0.739	0.032	0.049
CH_3	85% **P**(37%s + 62% p)+ 15% **Al**(12%s + 86% p)	76% **P**(23%s + 77%p) + 24% **Al**(22%s + 78%p)	0.489	0.728	0.036	0.052
c-C_3H_5	85% **P**(36%s + 64% p)+ 15% **Al**(13%s + 87% p)	76% **P**(22%s + 78%p) + 24% **Al**(22%s + 78%p)	0.493	0.726	0.037	0.051
C_6H_5	LP(**P**)→ 3p (**Al**)[a] 426	77% **P**(17%s + 83%p) + 23% **Al**(22%s + 78%p)	0.488	0.675	0.036	0.048

[a]The value reported corresponds to the second order interaction energy, in kJ mol^{-1}, between the P lone pair (LP(**P**)) and the empty 3p orbital of **Al** atom.

Table 16: Characteristics of the P–Ga bond in phosphine-gallanes in terms of atomic hybrids, Wiberg bond index, and electron density (ρ_b(P–Ga), a.u.) at the corresponding BCP.

	Bond		Wiberg bond index		$\rho_b(P-Ga)$	
R	neutral	anion	neutral	anion	neutral	anion
H	86% **P**(39%s + 61% p)+ 14% **Ga**(10%s + 89% p)	74% **P**(22%s + 78%p) + 26% **Ga**(22%s + 78%p)	0.491	0.787	0.048	0.068
CH_3	84% **P**(37%s + 63% p)+ 16% **Ga**(12%s + 88% p)	73% **P**(22%s + 78%p) + 27% **Ga**(22%s + 78%p)	0.541	0.784	0.055	0.072
c-C_3H_5	83% **P**(36%s + 64% p)+ 17% **Ga**(13%s + 87% p)	72% **P**(22%s + 78%p) + 28% **Ga**(23%s + 77%p)	0.562	0.799	0.056	0.071
C_6H_5	LP(**P**)→ 4p (**Ga**)[a] 510.5	74% **P**(17%s + 83%p) + 26% **Ga**(22%s + 78%p)	0.552	0.721	0.056	0.068

[a]The value reported corresponds to the second order interaction energy, in kJ mol^{-1}, between the P lone pair (LP(**P**)) and the empty 3p orbital of **Ga** atom.

Table 17: Calculated acidity ($\Delta_{acid}G^0$, kJ mol^{-1}) for R-XH_2 (R = Ethyl, vinyl, ethynyl; X = N, P, As, Sb) bases ($\Delta_r G_3^0$) and the corresponding R-XH_2·BeH_2, R-XH_2·BH_3, and R-XH_2·AlH_3 complexes ($\Delta_r G_4^0$).

	$\Delta_{acid}G^0$			
	$\Delta_r G_3^0$	$\Delta_r G_4^0$		
	free base	RXH_2·BeH_2	RXH_2·BH_3	RXH_2·AlH_3
R = Ethyl				
X = N	1627 (1638.9 ± 2.9)[a]	1431	1447	1399
X = P	1522 (1531. ± 12.)[b]	1409	1435	1360
X = As	1492 (1501. ± 8.8)[c]	1377	1393	1328
X = Sb	1455	1357	1361	1308
R = Vinyl				
X = N	1533	1325	1360	1322
X = P	1470 [1474][d] (1477. ± 9.6)[a]	1346 [1343][d]	1384	1340
X = As	1446 [1446][d] (1448. ± 8.8)[c]	1322 [1320][d]	1343	1309
X = Sb	1429	1316	1321	1295
R = Ethynyl				
X = N	1472	1282	1320	1283
X = P	1445 [1451][d] (1459. ± 9.6)[a]	1319 [1315][d]	1351	1305
X = As	1418 [1419][d] (1434. ± 8.8)[c]	1291 [1292][d]	1309	1277
X = Sb	1401 [1397][e]	1293	1290	1268

Table 18: Acidity enhancement, $\Delta\Delta_{acid}G^0$, and stabilization free energies of neutral, $\Delta_r G_1^0$, and deprotonated species, $\Delta_r G_2^0$, when R-XH_2 (R = Ethyl, vinyl, ethynyl; X = N, P, As, Sb) bases interact with BeH_2, BH_3, and AlH_3. All values are in kJ mol^{-1}.

	$(\Delta\Delta_{acid}G^0)^a$			$\Delta_r G_1^0$			$\Delta_r G_2^0$		
Y =	BeH_2	BH_3	AlH_3	BeH_2	BH_3	AlH_3	BeH_2	BH_3	AlH_3
R = Ethyl									
X = N	226	179	226	-59	-74	-66	-285	-253	-292
X = P	139	113	160	-12	-64	-26	-151	-177	-186
X = As	117	99	137	-17	-54	-34	-134	-153	-171
X = Sb	121	119	145	+17	-20	-26	-105	-139	-171
R = Vinyl									
X = N	208	171	209	-29	-41	-38	-237	-212	-247
X = P	124	78	129	-2[-11]^b	-65	-25	-126	-143	-154
X = As	124	102	136	+13[-2]^b	-24	-6	-111	-126	-142
X = Sb	114	106	132	+17	-15	-20	-97	-121	-153
R = Ethynyl									
X = N	190	152	188	-7	-16	-18	-197	-168	-206
X = P	125	93	138	+9[+1]^b	-43	-7	-116	-136	-145
X = As	127	106	140	+26[+9]^b	-11	+7	-101	-117	-133
X = Sb	108	110	132	+18[+17]^c	-7	-12	-90	-117	-144

[a] These values measure the acidity enhancement upon BeH_2, BH_3 and AlH_3 complexation and are given by the difference $\Delta_r G_3^0 - \Delta_r G_4^0$. [b] Values claculated at the G4 level fo theory. [c] Value calculated at the CCSD(T)/Def-QZVP level of theory.

Appendix 2

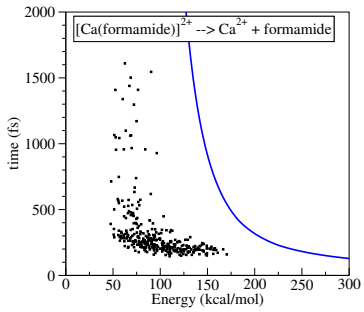


Figure 55: Reaction time *vs.* energy transfer for trajectories yielding formamide neutral loss, obtained from chemical dynamics simulations (squares) and half-life times ($t_{1/2}$) predicted by RRKM (solid lines). Both were obtained using BLYP/6-31G(d) level of theory.

Table 19: Relative errors (%) in A rotational constants. M = Ca.

Structure	B3LYP/6-31G(d)	G96LYP/6-31G(d)	MP2/6-31G(d)	BLYP/6-31G(d)
Formamide	-0.9	-2.4	-1.9	-2.8
A_$CaOH^+$	0.0	0.0	0.0	0.0
A_$HCNH^+$	0.0	0.0	0.0	0.0
B_$CaNH_3^{2+}$	-0.2	-1.4	0.1	-1.8
B_CO	0.0	0.0	0.0	0.0
C_$CaHCN^{2+}$	0.0	0.0	0.0	0.0
C_H_2O	-3.4	-6.7	-2.7	-7.6
D_CaH_2O^{2+}	-0.1	-1.4	-0.1	-1.7
D_HCN	0.0	0.0	0.0	0.0
E_$CaCO^{2+}$	0.0	0.0	0.0	0.0
E_NH_3	-1.9	-4.0	-1.0	-4.4
F_$CaOC^{2+}$	0.0	0.0	0.0	0.0
F_NH_3	-1.9	-4.0	-1.0	-4.4
G_$CaNH^{2+}$	-6.3	-13.6	-0.1	-16.3
G_OCH^+	0.0	0.0	0.0	0.0
int_10	1.3	2.8	-2.7	1.5
int_11	0.0	-1.3	0.2	-1.9
int_1	0.4	0.1	2.2	0.1
int_2	-0.5	-2.7	-0.2	-3.1
int_3	-1.3	-2.6	-1.7	-3.0
int_5	-0.3	-4.7	-3.6	-4.9
int_6	-0.6	-2.9	-1.5	-3.4
int_7	-14.0	-31.5	0.7	-33.2
int_8	2683.2	2673.1	2689.3	2642.1
int_9	-15.0	-10.6	-26.4	-11.8
TS_10_11	0.0	-2.7	6.1	-1.1
TS_1_10	-2.7	-9.6	-0.6	-9.5
TS_1_1	-1.6	-3.5	-2.3	-4.0
TS_1_2	0.0	-1.1	-1.6	-1.2
TS_1_5	1.8	-0.8	0.4	-1.3
TS_1_G	3.4	-5.0	29.4	-3.1
TS_2_3	-8.9	-13.7	-11.0	-14.4
TS_2_A	1.4	-0.8	1.4	-1.0
TS_3_6	-8.8	-11.6	-9.0	-12.4
TS_6_7	0.0	-2.7	-0.5	-2.0
TS_7_8	-3.1	-5.7	1.9	0.1
TS_8_9	4.5	-2.4	10.8	0.7
TS_9_A	18.5	33.7	21.7	67.0

Table 20: Relative errors (%) in B rotational constants. M = Ca.

Structure	B3LYP/6-31G(d)	G96LYP/6-31G(d)	MP2/6-31G(d)	BLYP/6-31G(d)
formamide	-0.91	-2.72	-1.25	-2.83
A_$CaOH^+$	-8.95	-9.7	-9.88	-10.08
A_$HCNH^+$	-1.47	-2.99	-3.26	-3.12
B_$CaNH_3^{2+}$	-6.19	-7.54	-7.28	-7.63
B_CO	-2.22	-4.18	-4.45	-4.34
C_$CaHCN^{2+}$	-5.25	-6.61	-7.61	-6.38
C_H_2O	-0.32	-1.25	-0.74	-1.21
D_CaH_2O^{2+}	-5.21	-6.51	-6.68	-6.26
D_HCN	-1.8	-3.58	-4.62	-3.7
E_$CaCO^{2+}$	-4.37	-5.75	-5.69	-5.2
E_NH_3	-1.77	-3.87	-0.92	-4.32
F_$CaOC^{2+}$	-3.28	-5.56	-5.99	-5.02
F_NH_3	-1.77	-3.87	-0.92	-4.32
G_$CaNH^{2+}$	-11.38	-14.1	-11.04	-15.25
G_OCH^+	-1.99	-3.82	-4.53	-4.02
int_10	-4.24	-6.33	-6.37	-4.71
int_11	0.0	-1.43	-0.88	-1.16
int_1	-4.6	-5.97	-5.68	-6.03
int_2	-3.14	-4.74	-4.12	-4.89
int_3	-4.51	-6.59	-5.27	-6.07
int_5	-6.26	-9.67	-6.06	-9.62
int_6	-4.92	-6.09	-6.11	-6.14
int_7	-4.19	-4.48	-7.88	-4.39
int_8	-24.3	-25.32	-25.7	-25.13
int_9	-3.77	-5.5	-4.45	-5.09
TS_10_11	0.0	-0.9	-7.16	0.02
TS_1_10	1.19	0.62	-1.04	-0.04
TS_1_1	-5.36	-6.73	-6.67	-6.67
TS_1_2	-3.99	-5.59	-4.86	-5.64
TS_1_5	-4.29	-5.75	-4.98	-5.58
TS_1_G	-13	-9.22	27.35	-11.89
TS_2_3	-0.58	-1.11	-0.94	-1.07
TS_2_A	-1.34	-5.15	-4.9	-5.68
TS_3_6	-3.02	-3.82	-3.82	-3.72
TS_6_7	-5.48	-6.86	-7.13	-7.08
TS_7_8	-2.0	-1.59	-11.71	-3.3
TS_8_9	-9.9	-7.35	-16.47	-8.75
TS_9_A	-22.14	-21.12	-21.72	-24.21

Table 21: Relative errors (%) in C rotational constants. M = Ca.

Structure	B3LYP/6-31G(d)	G96LYP/6-31G(d)	MP2/6-31G(d)	BLYP/6-31G(d)
formamide	-0,91	-2,66	-1,24	-2,76
A_$CaOH^+$	-8,95	-9,7	-9,88	-10,08
A_$HCNH^+$	-1,47	-2,99	-3,26	-3,12
B_$CaNH_3^{2+}$	-6,19	-7,54	-7,28	-7,63
B_CO	-2,22	-4,18	-4,45	-4,34
C_$CaHCN^{2+}$	-5,25	-6,61	-7,61	-6,38
C_H_2O	-1,42	-3,2	-1,44	-3,55
D_CaH_2O^{2+}	-5,12	-6,42	-6,57	-6,18
D_HCN	-1,8	-3,58	-4,62	-3,7
E_$CaCO^{2+}$	-4,37	-5,75	-5,69	-5,2
E_NH_3	0	-0,83	-0,29	-0,71
F_$CaOC^{2+}$	-3,28	-5,56	-5,99	-5,02
F_NH_3	0	-0,83	-0,29	-0,71
G_$CaNH^{2+}$	-11,03	-13,57	-10,8	-14,65
G_OCH^+	-1,99	-3,82	-4,53	-4,02
int_10	-3,86	-5,72	-6,13	-4,28
int_11	0	-1,44	-0,88	-1,16
int_1	-4,43	-5,77	-5,43	-5,83
int_2	-3,06	-4,68	-4,01	-4,84
int_3	-3,71	-5,61	-4,4	-5,31
int_5	-5,83	-9,34	-5,9	-9,3
int_6	-4,8	-6,01	-5,99	-6,07
int_7	-4,15	-4,37	-7,86	-4,27
int_8	-14,83	-15,98	-16,4	-15,77
int_9	-4,31	-5,73	-5,64	-5,4
TS_10_11	0	-1,12	-5,74	-0,12
TS_1_10	1,22	0,67	-1,04	0,03
TS_1_1	-4,47	-5,99	-5,66	-6,04
TS_1_2	-3,85	-5,43	-4,75	-5,49
TS_1_5	-4,15	-5,63	-4,86	-5,47
TS_1_G	-12,05	-8,6	27,6	-11,04
TS_2_3	-1,3	-2,32	-1,86	-2,37
TS_2_A	-1,32	-5,09	-4,78	-5,6
TS_3_6	-3,46	-4,36	-4,25	-4,34
TS_6_7	-5,2	-6,61	-6,8	-6,78
TS_7_8	-2,19	-2,32	-9,71	-2,74
TS_8_9	-8,35	-6,77	-13,86	-7,69
TS_9_A	-21,77	-20,68	-21,34	-23,64

Table 22: Relative errors (%) in A rotational constants. M = Sr.

Structure	B3LYP/6-31G(d)	G96LYP/6-31G(d)	MP2/6-31G(d)
formamide	2.12	0.57	1.03
A_HCNH$^+$	0.0	0.0	0.0
A_SrOH$^+$	0.0	0.0	0.0
B_CO	0.0	0.0	0.0
B_SrNH$_3^{2+}$	2.18	1.03	2.28
C_H_2O	-2.07	-5.21	-1.36
C_HCNSr^{2+}	-3.2	-7.89	-10.07
D_HCN	0.0	0.0	0.0
D_SrH_2O^{2+}	2.81	1.68	2.7
E_NH_3	-0.25	-2.31	0.6
E_OC-Sr^{2+}	0.0	0.0	0.0
F_CO-Sr^{2+}	0.0	0.0	0.0
F_NH_3	-0.25	-2.31	0.6
G_SrNH$_2^+$	-6.12	-11.18	2.38
G_OCH$^+$	0.0	0.0	0.0
int1	6.24	5.93	9.43
int2	2.82	0.54	3.78
int3	0.86	0.02	0.1
int4	2.92	0.97	2.79
int5	159.18	1184.09	4998.26
int6	-13.75	-11.99	-29.26
int7	20.38	17.63	14.11
int8	4.73	-4.43	2.57
TS_1_1	0.72	-1.38	-0.05
TS_1_2	2.55	1.62	0.7
TS_1_7	14.61	2.68	20.29
TS_1_8	3.12	-0.58	1.67
TS_2_3	-5.33	-11.56	-6.43
TS_2_A	19.29	18.02	20.94
TS_3_4	-4.49	-7.31	-6.33
TS_4_5	3.92	0.13	3.41
TS_5_6	26.97	15.26	28.7
TS_6_A	169.39	178.99	96.73
TS_1_G	13.24	2.73	268.41

Table 23: Relative errors (%) in B rotational constants. M = Sr.

Structure	B3LYP/6-31G(d)	G96LYP/6-31G(d)	MP2/6-31G(d)
formamide	1.74	-0.11	1.4
A_$HCNH^+$	1.57	0.02	-0.26
A_$SrOH^+$	-6.94	-8.19	-7.6
B_CO	1.9	-0.12	-0.42
B_$SrNH_3^{2+}$	-3.46	-4.8	-5.26
C_H_2O	3.04	2.0	2.6
C_$HCNSr^{2+}$	-3.56	-4.99	-6.42
D_HCN	1.87	0.03	-1.05
D_SrH_2O^{2+}	-3.37	-4.95	-5.08
E_NH_3	-0.25	-2.31	0.6
E_OC-Sr^{2+}	-2.99	-4.58	-5.02
F_CO-Sr^{2+}	-2.95	-6.21	-6.12
F_NH_3	-0.25	-2.31	0.6
G_$SrNH_2^+$	-10.07	-12.73	-10.82
G_OCH^+	1.85	-0.04	-0.78
int1	-3.57	-5.12	-5.06
int2	-1.63	-3.35	-3.25
int3	-3.06	-6.93	-3.68
int4	-2.53	-3.75	-4.56
int5	-18.99	-26.08	-27.67
int6	-3.03	-4.66	-2.35
int7	-7.1	-8.23	-9.41
int8	-5.68	-11.24	-5.9
TS_1_1	-3.67	-5.42	-5.73
TS_1_2	-2.28	-4.08	-3.47
TS_1_7	1.87	1.43	-0.28
TS_1_8	-1.15	-2.85	-2.74
TS_2_3	0.69	0.64	-0.76
TS_2_A	6.87	4.07	4.7
TS_3_4	-1.8	-2.75	-2.67
TS_4_5	-3.57	-4.95	-5.78
TS_5_6	-18.49	-16.01	-22.39
TS_6_A	-24.01	-23.38	-23.45
TS_1_G	-14.59	-10.47	-5.5

Table 24: Relative errors (%) in C rotational constants. M = Sr.

Structure	B3LYP/6-31G(d)	G96LYP/6-31G(d)	MP2/6-31G(d)
formamide	1.8	0.0	1.46
A_HCNH$^+$	1.57	0.02	-0.26
A_SrOH$^+$	-6.94	-8.19	-7.6
B_CO	1.9	-0.12	-0.42
B_SrNH$_3^{2+}$	-3.46	-4.8	-5.26
C_H$_2$O	1.23	-0.59	1.2
C_HCNSr^{2+}	-3.56	-4.99	-6.42
D_HCN	1.87	0.03	-1.05
D_SrH$_2$O^{2+}	-3.3	-4.87	-4.99
E_NH$_3$	3.15	2.4	2.84
E_OC-Sr^{2+}	-2.99	-4.58	-5.02
F_CO-Sr^{2+}	-2.95	-6.21	-6.12
F_NH$_3$	3.15	2.4	2.84
G_SrNH$_2^+$	-9.75	-12.32	-10.61
G_OCH$^+$	1.85	-0.04	-0.78
int1	-3.34	-4.87	-4.74
int2	-1.54	-3.27	-3.11
int3	-2.41	-5.81	-3.05
int4	-2.44	-3.67	-4.43
int5	-7.77	-11.55	-13.2
int6	-3.47	-4.95	-3.7
int7	-5.5	-6.71	-8.01
int8	-5.16	-10.94	-5.48
TS_1_1	-2.96	-4.77	-4.82
TS_1_2	-2.17	-3.94	-3.37
TS_1_7	1.79	1.44	-0.38
TS_1_8	-1.09	-2.82	-2.7
TS_2_3	0.27	-0.26	-1.25
TS_2_A	6.43	3.68	4.35
TS_3_4	-2.04	-3.04	-3.01
TS_4_5	-3.32	-4.74	-5.48
TS_5_6	-14.96	-13.28	-18.65
TS_6_A	-23.25	-22.61	-22.81
TS_1_G	-13.76	-9.96	-2.51

Table 25: Absolute (kcal mol^{-1}) and relative errors (%) for relative energies computed with the four trial methods. M = Ca.

	Absolute error (kcal mol^{-1})				Relative error(%)			
Structure	B3LYP[a]	G96LYP[a]	MP2[a]	BLYP[a]	B3LYP[a]	G96LYP[a]	MP2[a]	BLYP[a]
Form/Ca^{2+}	6.59	11.15	9.23	7.22	-6.9	-11.7	-9.7	-7.6
A	25.29	21.07	17.28	20.89	-528.6	-440.5	-361.2	-436.6
B	1.28	4.97	10.51	2.64	-3.1	-12.1	-25.5	-6.4
C	2.35	1.54	9.93	0.73	4.2	-2.7	-17.6	1.3
D	1.64	2.71	12.32	0.38	2.7	-4.5	-20.4	-0.6
E	1.16	6.22	11.60	3.91	-1.5	-8.0	-15.0	-5.0
F	2.51	5.65	9.43	3.22	-3.1	-6.9	-11.5	-3.9
G	19.41	12.64	13.67	11.43	58.4	38.0	41.1	34.4
int10	2.51	4.97	11.50	4.93	-7.3	-14.5	-33.6	-14.4
int11	4.03	6.15	13.79	6.67	-21.1	-32.2	-72.1	-34.8
min1	0.00	0.00	0.00	0.00	0.0	0.0	0.0	0.0
int2	2.00	3.75	4.72	4.50	-3.8	-7.1	-8.9	-8.5
int3	2.49	2.06	5.05	2.75	-12.1	-10.0	-24.6	-13.3
int4	0.72	2.53	2.41	3.33	-1.3	-4.5	4.3	-6.0
int5	0.94	3.04	2.59	3.99	-2.0	-6.5	-5.5	-8.5
int6	0.45	1.59	0.16	1.63	-1.6	-5.6	-0.6	-5.8
int7	4.33	1.93	5.30	2.25	14.7	6.5	-18.0	7.6
int8	1.65	0.81	10.63	0.36	31.3	15.3	-201.9	6.9
int9	4.41	2.02	7.20	1.84	16.6	7.6	-27.1	6.9
TS_10_11	4.20	6.56	13.48	6.58	-11.6	-18.2	-37.3	-18.2
TS_1_10	4.95	1.23	2.28	2.90	6.5	-1.6	3.0	-3.8
TS_1_1	4.85	4.55	7.04	5.51	-16.7	-15.7	-24.3	-19.0
TS_1_2	0.86	4.72	3.61	4.93	-1.4	-7.5	-5.7	-7.8
TS_1_5	1.10	3.72	3.08	4.23	1.2	-4.1	3.4	-4.6
TS_1_G	12.67	2.31	7.72	1.61	15.9	2.9	9.7	2.0
TS_2_3	1.71	3.22	4.33	3.86	-3.1	-5.9	-7.9	-7.0
TS_2_A	5.06	0.89	1.56	0.34	9.2	1.6	2.8	-0.6
TS_3_6	2.41	1.06	5.02	0.76	4.5	2.0	9.5	1.4
TS_6_7	2.14	2.59	0.40	2.97	2.7	-3.3	0.5	-3.8
TS_7_8	4.00	2.23	7.61	1.99	10.4	5.8	-19.8	5.2
TS_8_9	3.34	1.45	8.66	1.13	11.8	5.1	-30.7	4.0
TS_9_A	19.23	15.20	11.01	14.65	47.0	37.2	26.9	35.8

[a] 6-31G(d) basis set.

Table 26: Absolute (kcal mol^{-1}) and relative errors (%) for relative energies computed with the four trial methods. M = Sr.

	Absolute error (kcal mol^{-1})				Relative error(%)			
Structure	B3LYP[a]	G96LYP[a]	MP2[a]	G96LYP[b]	B3LYP[a]	G96LYP[a]	MP2[a]	G96LYP[b]
Form/Sr^{2+}	13.25	18.74	16.41	6.14	-16.7	-23.6	-20.7	-7.7
A	45.16	40.10	37.33	7.14	-755.1	-670.4	-624.2	-119.4
B	4.80	9.05	14.61	2.29	-12.9	-24.4	-39.4	-6.2
C	4.93	0.76	8.09	0.78	10.8	1.7	-17.7	1.7
D	0.33	5.21	15.29	2.74	-0.6	-9.6	-28.3	-5.1
E	2.23	7.67	12.44	0.02	-3.4	-11.6	-18.8	0.0
F	4.32	8.00	11.97	1.44	-6.1	-11.2	-16.8	-2.0
G	33.58	25.21	31.59	9.29	99.9	75.0	94.0	27.6
min1	0.00	0.00	0.00	0.00	0.0	0.0	0.0	0.0
int2	3.96	5.50	6.26	1.03	-8.0	-11.1	-12.6	-2.1
int3	5.45	5.20	6.66	1.07	-24.5	-23.4	-29.9	-4.8
int6	2.97	4.25	2.11	0.01	-10.8	-15.4	-7.7	0.0
int8	1.23	0.18	11.04	0.81	13.2	1.9	-118.3	8.7
int9	3.58	1.42	8.27	0.33	14.4	5.7	-33.2	1.3
int10	5.11	7.74	13.22	0.00	-16.0	-24.3	-41.5	0.0
int5	0.56	2.57	0.09	1.16	1.2	-5.7	0.2	2.6
TS_1_1	6.23	5.86	7.45	0.03	-21.0	-19.8	-25.1	-0.1
TS_1_2	0.32	4.23	2.46	1.74	-0.5	-7.2	-4.2	-2.9
TS_1_10	9.64	2.83	9.81	3.20	13.0	3.8	13.3	4.3
TS_1_5	4.38	0.81	6.75	0.41	5.0	-0.9	7.8	-0.5
TS_2_3	3.89	5.12	6.06	1.09	-7.5	-9.9	-11.7	-2.1
TS_2_A	6.99	3.97	5.59	2.01	13.6	7.7	10.9	3.9
TS_3_6	2.00	0.37	5.54	0.92	3.8	0.7	10.6	1.8
TS_6_8	3.02	1.94	1.73	0.96	4.0	-2.5	2.3	-1.3
TS_8_9	2.15	0.21	9.95	0.56	8.1	0.8	-37.5	2.1
TS_9_A	35.87	31.39	28.58	5.49	87.7	76.7	69.9	13.4
TS_1_G	25.91	11.99	26.88	5.59	34.1	15.8	35.4	7.3

[a] 6-31G(d) basis set. [b] 6-31+G(d,p) basis set.

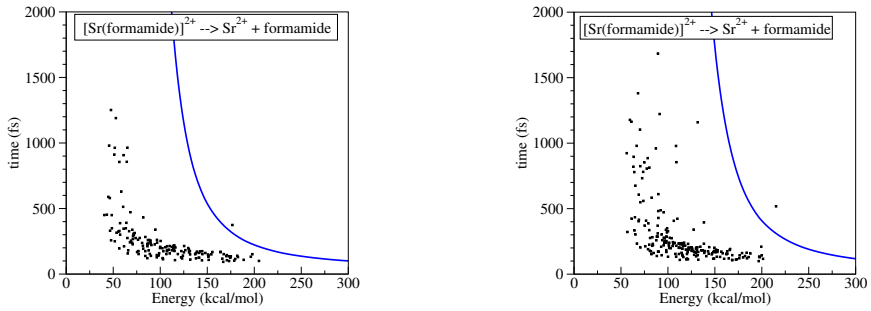


Figure 56: Reaction time *vs.* energy transfer obtained from chemical dynamics simulations (squares) and half-life times ($t_{1/2}$) predicted by RRKM (solid lines). Both are for trajectories yielding formamide neutral loss using G96LYP/6-31G(d) in the left and G96LYP/6-31+G(d,p) in the right.